```
GTCGACCCACGCGTCCGGGCGGCCCGCGGCAGTGCAGGCTGCAGGCGGCCGAGGAGGCTGCCGCTCTGGCTTGCCG          79

                                                 M   D   L   P   R   G   L   V            8
CCCCCCGCGGCCGCTGCACACCGGACCCAGCCGCCGGTCCGGCGCC  ATG GAC CTG CCC AGG GGC CTG GTG        150

 V   A   W   A   L   S   L   W   P   G   F   T   D   T   F   N   M   D   T   R            28
GTG GCC TGG GCG CTC AGC CTG TGG CCA GGG TTC ACG GAC ACC TTC AAC ATG GAC ACC AGG        210

 K   P   R   V   I   P   G   S   R   T   A   F   F   Y   G   Y   T   V   Q   H            48
AAG CCC CGG GTC ATC CCT GGC TCC AGG ACC GCC TTC TTT TAC GGC TAC ACA GTG CAG CAC        270

 D   I   S   G   N   K   W   L   V   V   G   A   P   L   E   T   N   G   Y   Q            68
GAC ATC AGT GGC AAT AAG TGG CTG GTC GTG GGC GCC CCA CTG GAA ACC AAT GGC TAC CAG        330

 K   T   G   D   V   Y   K   C   P   V   I   H   G   N   C   T   K   L   N   L            88
AAG ACG GGA GAC GTG TAC AAG TGT CCA GTG ATC CAC GGG AAC TGC ACC AAA CTC AAC CTG        390

 G   R   V   T   L   S   N   V   S   E   R   K   D   N   M   R   L   G   L   S           108
GGA AGG GTC ACC CTG TCC AAC GTG TCC GAG CGG AAA GAC AAC ATG CGC CTC GGC CTT AGT        450

 L   A   T   N   P   K   D   N   F   L   A   C   S   P   L   W   S   H   E               128
CTC GCC ACC AAC CCC AAG GAC AAC TTC CTG GCC TGC AGC CCC CTC TGG TCT CAT GAG            510

 C   G   S   Y   Y   T   T   G   M   C   S   R   V   N   S   N   F   R   F               148
TGT GGG AGC TAC TAC ACC ACA GGG ATG TGT TCA AGA GTC AAC TCC AAC TTC AGG TTC            570

 S   K   T   V   A   P   A   L   Q   R   C   Q   T   Y   M   D   I   V                   168
TCC AAG ACC GTG GCC CCA GCT CTC CAA AGG TGC CAG ACC TAC ATG GAC ATC GTC                630

 L   D   G   S   N   S   I   Y   P   W   V   E   V   Q   H   F   L   I   N   I           188
CTG GAT GGC TCC AAC AGC ATC TAC CCC TGG GTG GAG GTT CAG CAC TTC CTC ATC AAC ATC        690

 L   K   K   F   Y   I   G   P   G   Q   I   Q   V   G   V   V   Q   Y   G   E           208
CTG AAA AAG TTT TAC ATT GGC CCA GGG CAG ATC CAG GTT GGA GTT GTG CAG TAT GGC GAA        750
```

FIG. 1A

```
D   V   V   H   E   F   H   L   N   D   Y   R   S   V   K   D   V   V   E   A      228
GAT GTG GTG CAT GAG TTT CAC CTC AAC GAC TAC AGG TCT GTA AAA GAT GTG GTG GAA GCT    810

A   S   H   I   E   Q   R   G   G   T   E   T   R   T   A   F   G   I   E   F      248
GCC AGC CAC ATT GAG CAG AGA GGA GGA ACA GAG ACC CGG ACG GCA TTT GGC ATT GAA TTT    870

A   R   S   E   A   F   Q   K   G   R   K   G   A   K   K   V   M   I   V          268
GCA CGC TCA GAG GCT TTC CAG AAG GGT GGA AGG AAA GGA GCC AAG GTG ATG ATT GTC        930

I   T   D   G   E   S   H   D   S   D   L   E   K   V   I   Q   S   E              288
ATC ACA GAT GGG GAG TCC CAC GAC AGC GAC CTG GAG AAG GTG ATC CAA AGC GAA            990

R   D   N   V   T   R   Y   A   V   A   L   G   Y   L   N   R   R   G   I          308
AGA GAC AAC GTA ACA AGA TAT GCG GTG GCC CTG GGC TAC CTG AAC CGC AGG GGG ATC        1050

N   P   E   T   F   L   N   E   I   K   Y   I   A   S   D   P   D   D   K   H      328
AAT CCA GAA ACT TTT CTA AAT GAA ATC AAA TAC ATC GCC AGT GAC CCT GAT GAC AAG CAC    1110

F   N   V   T   D   E   A   A   L   K   D   I   V   D   A   L   G   D   R          348
TTC AAT GTC ACT GAT GAG GCT GCC TTG AAG GAC ATT GTC GAT GCC CTG GGG GAC AGA        1170

I   F   S   L   E   G   T   N   K   N   E   T   S   F   G   L   E   M   S   Q      368
ATC TTC AGC CTG GAA GGC ACC AAC AAG AAC GAG ACC TCC TTT GGG CTG GAG ATG TCA CAG    1230
```

FIG. 1B

```
  T   G   F   S   S   H   V   V   E   D   G   V   L   L   G   A   V   G   A   Y    388
  ACG GGC TTT TCC TCG CAC GTG GTG GAG GAT GGG GTT CTG CTG GGA GCC GTC GGT GCC TAT  1290

  D   W   N   G   A   V   L   K   E   T   S   A   G   K   V   I   P   L   R   E    408
  GAC TGG AAT GGA GCT GTG CTA AAG GAG ACG AGT GCC GGG AAG GTC ATT CCT CTC CGC GAG  1350

  S   Y   L   K   E   F   P   E   E   P   A   Y   G   A   L   G   N   H   G   Y    428
  TCC TAC CTG AAA GAG TTC CCC GAG GAG CCC GCA TAC GGT GCA CTG GGT AAC CAT GGA TAC  1410

  V   T   S   V   S   R   Q   S   R   G   R   V   Y   V   A   G   V   A   R   F    448
  GTC ACA TCG GTG TCC AGG CAG TCC CGG GGG CGA GTG TAC GTG GCC GGA GTC GCC CGG TTC  1470

  N   H   T   G   K   V   I   L   F   T   M   H   N   N   R   S   L   T   I   H    468
  AAC CAC ACG GGC AAG GTC ATC CTG TTC ACC ATG CAC AAC AAC CGG AGC CTC ACC ATC CAC  1530

  Q   A   M   R   G   Q   Q   I   G   S   Y   F   G   S   E   I   T   S   V   D    488
  CAG GCT ATG CGG GGC CAG CAG ATA GGC TCT TAC TTT GGG AGT GAA ATC ACC TCG GTG GAC  1590

  I   D   G   D   G   V   T   D   V   L   V   L   G   A   P   M   Y   F   N   E    508
  ATC GAC GGC GAC GGC GTG ACT GAT GTC CTG GTG CTG GGC GCA CCC ATG TAC TTC AAC GAG  1650

  G   R   E   R   G   K   V   Y   V   Y   E   L   R   Q   N   R   F   V   Y   N    528
  GGC CGT GAG CGA GGC AAG GTG TAC GTC TAT GAG CTG AGA CAG AAC CGG TTT GTT TAT AAC  1710

  G   T   L   K   D   S   H   Y   S   Q   N   A   R   F   G   S   I   A   S        548
  GGA ACG CTA AAG GAT TCA CAC TAC AGT CAG AAT GCC CGA TTT GGG TCC ATT GCC TCA      1770

  V   R   D   L   N   Q   D   S   Y   N   D   V   V   G   A   P   L   E   D        568
  GTT CGA GAC CTC AAC CAG GAT TCC TAC AAT GAC GTG GTG GGA GCC CCC CTG GAG GAC      1830

  N   H   A   G   A   I   Y   F   H   G   F   R   G   S   I   L   K   T   P        588
  AAC CAC GCA GGA GCC ATC TAC TTC CAC GGC TTC CGA GGC AGC ATC CTG AAG ACA CCT      1890

  K   Q   R   I   T   A   S   E   L   A   T   G   L   Q   Y   F   G   C   S   I    608
  AAG CAG AGA ATC ACA GCC TCA GAG CTG GCT ACC GGC CTC CAG TAT TTT GGC TGC AGC ATC  1950
```

FIG. 1C

```
                                                                                628
H   G   Q   L   D   L   N   E   D   G   L   I   D   L   A   V   G   A   L   G    2010
CAC GGG CAA TTG GAC CTC AAT GAG GAT GGG CTC ATC GAC CTG GCA GTG GGA GCC CTT GGC

                                                                                648
N   A   V   I   L   W   S   R   P   V   Q   I   N   A   S   L   H   F   E        2070
AAC GCT GTG ATT CTG TGG TCC CGC CCA GTG CAG ATC AAT GCC AGC CTC CAC TTT GAG

                                                                                668
P   K   S   I   N   I   F   H   R   D   C   K   R   S   G   R   D   A   T   C    2130
CCA AAG TCC ATC AAC ATC TTC CAC AGA GAC TGC AAG CGC AGT GGC CGC GAT GCC ACC TGC

                                                                                688
L   A   F   L   C   F   L   P   A   L   H   F   Q   H   P   A   T   T   T        2190
CTG GCC TTC CTC TGC TTC CTC CCC GCA CTG CAT TTC CAA CCC GCA ACA ACA ACT

                                                                                708
V   G   I   R   Y   N   A   T   M   D   E   R   Y   T   P   R   A   H   L        2250
GTT GGC ATC AGA TAC AAC GCC ACC ATG GAT GAG AGG CGG TAT ACA CCG AGG GCC CAC CTG

                                                                                728
D   E   G   G   D   R   F   T   N   R   A   V   L   S   S   G   E   L            2310
GAC GAG GGG GGC GAC CGA TTC ACC AAC AGA GCC GTA CTG TCC TCC GGC CAG GAG CTC

                                                                                748
C   E   R   I   N   F   H   V   L   D   T   A   D   Y   V   K   P   V   T   F    2370
TGT GAG CGG ATC AAC TTC CAT GTC CTG GAC ACT GCT GAC TAC GTG AAG CCA GTG ACC TTC

                                                                                768
S   V   E   Y   S   L   E   D   P   H   G   P   M   L   D   G   W   P            2430
TCA GTC GAG TAT TCC CTG GAG GAC CCT CAT GGC CCC ATG CTG GAC GGC TGG CCC
```

FIG. 1D

```
      T   T   L   R   V   S   V   P   F   W   N   G   C   N   E   D   E   H   C   V    788
     ACC ACT CTC AGA GTC TCG GTG CCC TTC TGG AAC GGC TGC AAT GAG GAT GAG CAC TGT GTC  2490

      P   D   L   V   L   D   A   R   S   D   L   P   T   A   M   E   Y   C   Q   R    808
     CCT GAC CTT GTG TTG GAT GCC CGG AGT GAC CTG CCC ACG GCC ATG GAG TAC TGC CAG AGG  2550

      V   L   R   K   P   A   Q   D   C   S   A   Y   L   S   F   D   T   V         828
     GTG CTG AGG AAG CCT GCG CAG GAC TGC TCC TAC CTG TCC TTC GAC ACC ACA GTC          2610

      F   I   E   S   T   R   Q   V   A   V   E   A   T   L   E   N   R   G         848
     TTC ATA GAG AGC ACA CGC CAG CGA GTG GCG GAG GCC ACA CTG GAG AAC AGG GGC          2670

      E   N   A   Y   S   T   V   L   N   I   S   Q   A   N   L   Q   F   A   S    868
     GAG AAC GCC TAC AGC ACG GTC CTA AAT ATC TCG CAG TCA AAC CTG CAG TTT GCC AGC     2730

      L   I   Q   K   E   D   S   D   G   S   I   E   C   V   N   E   E   R   L    888
     TTG ATC CAG AAG GAG GAC TCA GAC GGT AGC ATT GAG TGT GTG AAC GAG GAG AGG CTC     2790

      Q   K   Q   V   C   N   V   S   Y   P   F   F   R   A   K   A   K   V   A   F  908
     CAG AAG CAA GTC TGC AAC GTC AGC TAT CCC TTC TTC CGG GCC AAG GCC AAG GTG GCT TTC  2850

      R   L   D   F   E   F   S   K   S   I   F   L   H   H   L   E   I   E   L   A  928
     CGT CTT GAT TTT GAG TTC AGC AAA TCC ATC TTC CTA CAC CAC CTG GAG ATC GAG CTC GCT  2910

      A   G   S   D   N   E   R   D   S   T   K   E   D   N   V   A   P   L   R    948
     GCA GGC AGT GAC AAT GAG CGG GAC AGC ACC AAG GAA GAC AAC GTG GCC CCC TTA CGC     2970

      F   H   L   K   Y   E   A   D   V   L   F   T   R   S   S   L   S   H   Y    968
     TTC CAC CTC AAA TAC GAG GCT GAC GTC CTC TTC ACC AGG AGC AGC CTG AGC CAC TAC     3030

      E   V   K   P   N   S   L   E   R   Y   D   G   I   G   P   P   F   S   C    988
     GAG GTC AAG CCC AAC AGC TCG CTG GAG AGA TAC GAT GGT ATC GGG CCT CCC TTC AGC TGC  3090

      I   F   R   I   Q   N   L   G   L   F   P   I   H   G   M   M   K   I   T   1008
     ATC TTC AGG ATC CAG AAC TTG GGC TTG TTC CCC ATC CAC GGG ATG ATG AAG ATC ACC     3150
```

FIG. 1E

```
 I   P   I   A   T   R   S   G   N   R   L   L   K   L   R   D   F   L   T   D    1028
ATT CCC ATC GCC ACC AGG AGC GGC AAC CGC CTA CTG AAG CTG AGG GAC TTC CTC ACG GAC  3210

 E   A   N   T   S   C   N   I   W   G   N   S   T   E   Y   R   D   P   T   V    1048
GAG GCG AAC ACG TCC TGT AAC ATC TGG GGC AAT AGC ACT GAG TAC CGG GAT CCA ACC GTG  3270

 E   D   L   R   R   A   P   Q   L   N   H   S   N   D   V   V   S   I            1068
GAG GAC TTG CGT CGT GCT CCA CAG CTG AAT CAC AGC AAC GAT GTC GTC TCC ATC          3330

 N   C   N   I   R   L   V   P   N   Q   E   I   N   F   H   L   L   G   N   L    1088
AAC TGC AAT ATA CGG CTA GTC CCC AAC CAG GAA ATC AAT TTC CAT CTA CTG GGG AAC CTG  3390

 W   L   R   S   L   K   A   L   K   Y   K   S   M   I   V   N   A   A            1108
TGG TTG AGG TCC CTA AAA GCA CTC AAG TAC AAA TCC ATG ATC GTC AAC GCA GCC          3450

 L   Q   R   Q   F   H   S   P   F   I   F   R   E   E   D   P   S   R   Q   I    1128
TTG CAG AGG CAG TTC CAC AGC CCC TTC ATC TTC CGT GAG GAG GAT CCC AGC CGC CAG ATC  3510

 V   F   E   I   S   K   Q   E   D   W   Q   V   P   I   W   I   V   G   S        1148
GTG TTT GAG ATC TCC AAG CAA GAG GAC TGG CAG GTC CCC ATC TGG ATC GTA GGC AGC      3570

 T   L   G   G   L   L   L   A   L   L   V   L   K   L   G   F                    1168
ACC CTG GGG GGC CTC CTA CTG GCA CTG CTG GTC CTG AAG CTC GGC TTC                  3630

 F   R   S   A   R   R   R   E   P   G   L   D   P   T   P   K   V   L   E        1188
TTT AGA AGT GCC CGC AGG AGG GAG CCT GGT CTG GAC CCC ACC CCC AAA GTG CTG GAG      3690

 *                                                                               1189
TGA                                                                              3693
```

FIG. 1F

GGCTCCAGAGGAGGAGACTTTGAGTTGATGGGGGCCAGGACCAGTCCAGTAGTGTTGAGACCCAGCCTGTGGCCCCAC 3772

CGAGCTGGAGCGGAGAGGAAGCCAGCTGGCTTTGCACTTGACCTCATCTCCCGAGCAATGGCGCGCTGCCTCCAGAA 3851

TGGAACTCAAGCTGGTTTTAAGTGGAACTGCCCTACTGGGACACCTTTAACACAGACCCCTAGGGATTTAA 3930

AGGGACACCCCTACACACCCAGGCCCACGCCCAAGGCCTCCCTCGTGTGGAGGCATTGCTGCCCCAGCTAC 4009

TAAGGTGCTAGGAATTCGTAATCATCCCCATCCTCCAGAGAAACCCAGGGAGGAAGACTGTAAATACGAACCCAATCTG 4088

CACACTCCAGGCCTCTAGTTCCAGAAGGATCCAAGACAAAACAGATCTGAATTCTGCCCTTTCTCTCACCCATCCCAC 4167

CCCTCCATTGGCTCCCAAGTCACACCCACTCCCTTCCCCATAGATAGGCCCCTGGGGCTCCTGAAGAATGAACCCAAGA 4246

GCAAGGGCTTGATGTGTGACAGCTGCAAGCCCAGGGATGAAGAAAGACTCTGAGATGTGGAGACTGATGGCCAGGCAAGTG 4325

GGACCAGGATACTGGACGCTGTCCTGAGATGAGGAGGTAGCCGGGCTCTGCACCCAGTGCATTCACATTGACCGCAACT 4404

CACACAATTCCCCACCAGCTGCAGCCCCTTGCTCTCAGCTGCCAACCCTCCCGGGTCACTTTGTTCCCAGTACCTCA 4483

TGGGAAGCATGTGATGACAATCCCTGGGGCTGTGCATTCCCACGTCTCTTCTGCTGCCCCTCCCCTGTATCACCCCCTACACA 4562

CGCACTGGCCTGGCAGCTGGGCAGCTGGGGTAGGGGTAGGAGCCTCCCCTCATGCCCCAGTGCACAGGGAAGGGCTTGGCCAGC 4641

CACACACACACACACACACCACTGCCTCCCATCCTCCCTGGAATGCACTGAATAAAGCACGTGCAAGGACTCCCGGAGCCTTGGTGCA 4720

GCTGTTGAGGGGTCCCCTCTGGAATGCACTGAATAAAGCACGTGCAAGGACTCCCGGAGCCTTGGTGCA 4799

AATATCTCATCTGCCGGCCCCCAGGACAAGTGTATGACCCAGTGATAATGCCCAGGACAAGGGCGTGCCTGGCGCC 4878

CAGTGGAGTAATTTATGCCTTAGTCTTGTTTGTTTGAGGTAGAAATGCAAGGGGGACACATGAAAGGCATCAGTCCCCCTGT 4957

GCATAGTACGACCTTTACTGTCGTATTTTGAAAAATTAAAAATACAGTGTTTAAAAACAAAAAAAAAAAAAAGGGC 5036
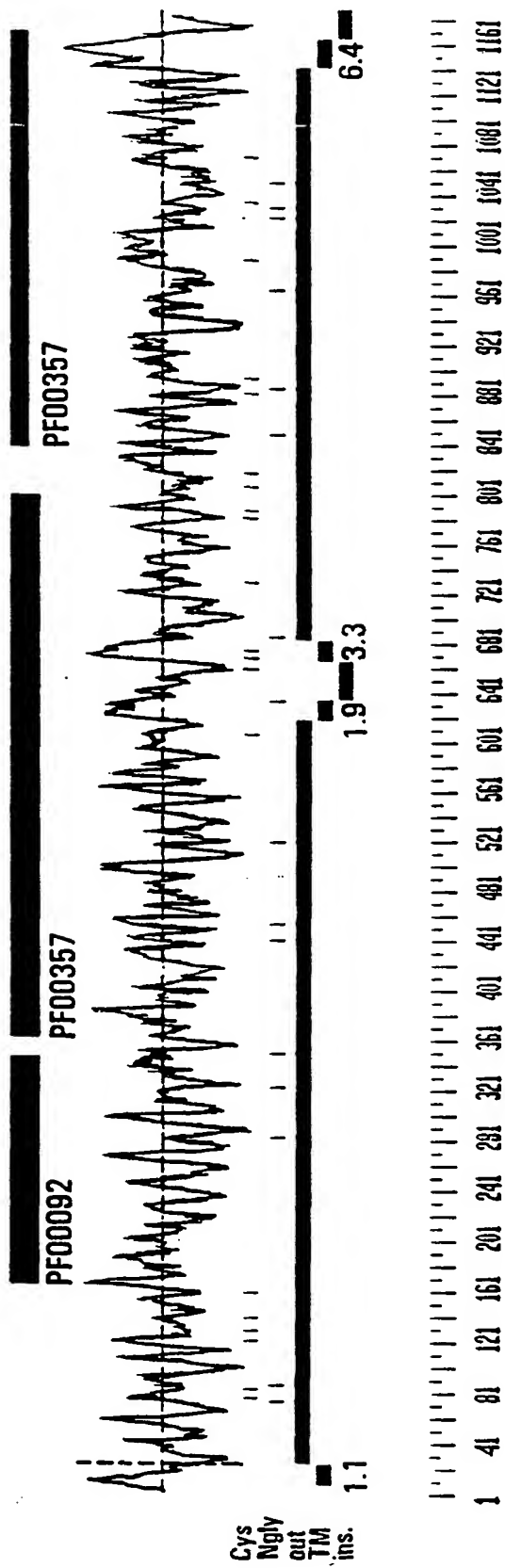
GGCCGC 5042

PF00357

PF00357

PF00092

Cys
Ngly
out
TM
ins.

1.1   1.9 3.3   6.4

1 41 81 121 161 201 241 281 321 361 401 441 481 521 561 601 641 681 721 761 801 841 881 921 961 1001 1041 1081 1121 1161

FIG. 2

```
              10        20        30        40        50        60        70
inputs GTCGACCCACGCGTCCGGGCGCCCCGCGCAGTGCAGGCTGCAGGCGCCGCGCCGAGGAGGCTGCCGCTCT
                                   ::: :::.                              ..::: ::.
       ---------------------------CAGGTCAGA------------------------AACCGATCA
                                                                       10

              80        90        100       110       120       130       140
inputs GGCTTGCCGCCCCCGCCGCCGCTGCACACCGGACCCAGCCGCCGTGCCGCGGGCCATGGACCTGCCCAG
       \:::.:: .: ::::: :: :.:: :::: :.  : .::: ::  :    :: .::::: :
       GGCATGGAACTCCCCTTCGTCACT---CACCTGTTCTTGCC-CCTGG----------TGTTCCTGAC---
        20        30        40         50        60                70

              150       160       170       180       190       200       210
inputs GGGCCTGGTGGTGGCCTGGGCGCTCAGCCTGTGGCCAGGGTTCACGGACACCTTCAACATGGACACCAGG
        .:: ::    :: : : .: :    ::: :..:.::.: ::. ::       :::..
       AGGTCT--------CT--GCTCCC--CCTTTAACCTGGATGAACA-TCACC--CA------CGCCTAT
                 80         90        100       110

              220       230       240       250       260       270       280
inputs AAGCCCCGGGTCATCCCTGGCTCCAGGACCGCCTTCTTTGGCTACACAGTGCAGCAGCACGACATCAGTG
        .   :::.:: :        :.:::::.: ::  : :::::  ::::  .::  .-:::.:: :. .  .:::
       T---CCCAGGGC--------CACCAGAA--GCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTG
        120              130       140       150       160       170

              290       300       310       320       330       340
inputs GCAATAAGTGGCTGGTCGTGGGCGCCCCACTGGAAA---CCAA-TGGCTACCAGAAGACGGGAGACGTGT
        :   :  ..:::: :: : :::::::::::: :::::.:.   ::.. .::: ::::::.:   ::.:::::: :
       GACAGCGATGGATGCTGGTGGGCGCCCC-CTGGGATGGGCCTTCAGGCGACCGGAGG---GGGGACGTTT
        180       190       200       210       220       230        240

          350       360       370       380       390       400
inputs ACAAGTGTCCAGT------GATCCACGGGAAC---TGCACCAAACTCAACCTGGGAAGGGTCACCCTGTC
        :  .: :::    : .:  : ::  :.. . :    :: .::::: .  . :::
       ATCGCTGCCCTGTAGGGGGGGCCCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGG
            250       260       270        280       290       300      310

          410       420       430       440       450       460       470
inputs CAACGTGTCCGAGCGGAAAGACAACATGCGCCTCGGCCTTAGTCTCGCCACCAACCCCAAGG--ACAACA
        ::   .:: : :  .:. :: :: :::::.::: ::   .:.::::   :. : :.:: :  .. .
       AAATTCATCTCATCCTGCTGTGAATATGCACCTGGG--GATGTCTCTGTTAGAGACAGATGGTGATGGGG
        320       330       340        350       360       370

          480       490       500       510       520       530       540
inputs GCTTCCTGGCCTGCAGCCCC-CTCTGGTCTCATGAGTGTGGGAGCTCCTACTACACCACAGGGATGTGTT
        : ::: :::::::: .::::: ::::::::::::..:: ::::: ::::: ..:.:: ..:.:::::.:::
       GATTCATGGCCTG-TGCCCCTCTCTGGTCTCGTGCTTGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTG
        380       390       400       410       420       430       440

          550       560       570       580       590       600       610
inputs C--AAGAGTCAACTCCAACTTCAGGTTCTCCAAGACCGTGGCCCCAGCTCTCCAAAGGTGCCAGACCTAC
        :  ...: .: ::: ::. : ::: ::. ....:.  :: :::: : :::: : :::: ::: :::
       CCCGTGTGGATGCTTCAT-TCCAGCCTCAGGGAAGCC-TGGCACCCACTGCCCAACGCTGCCCAACATAC
       450       460       470       480       490       500       510

          620       630       640       650       660       670       680
inputs ATGGACATCGTCATTGTCCTGGATGGCTCCAACAGCATCTACCCCTGGGTGGAGGTTCAGCACTTCCT-C
        :::: .: :: ::::::::: :::::::::::::::::::::::::::  ::.:::::: ::::: :
       ATGGATGTTGTCATTGTCTTGGATGGCTCCAACAGCATCTACCCCTGGTCTGAAGTTCAGACCTTCCTAC
       520       530       540       550       560       570       580
```

FIG. 3A

```
             690       700       710       720       730       740       750
inputs ATCAACATCCTGAAAAAGTTTTACATTGGCCCAGGGCAGATCCAGGTTGGAGTTGTGCAGTATGGCGAAG
       .. ..:.  .:...:: : :.. ::::.:.::::.:::: :::.: ::. : :..:.:::::: ::..
       GAAGACTGGTAGGGAAACTGTTT-ATTGACCCAGAACAGATACAGGTGGGACTGGTACAGTATGGGGAGA
             590       600       610       620       630       640       650


             760       770       780       790       800       810       820
inputs ATGTGGTGCATGAGTTTCACCTCAACGACTACAGGTCTGTAAAAGATGTGGTGGAAGCTGCCAGCCACAT
       ::.::::::: :   ::: .. :: :.: ::.: ::::::.::::::::..::.. :: :.  :: :
       GCCCTGTACATGAGTGGTCCCTGGGAGATTTCCGAACGAAGGAAGAAGTGGTGAGAGCAGCAAAGAACCT
             660       670       680       690       700       710       720


             830       840       850       860       870       880
inputs TGAGCAGAGAGGAGGAACAGAGACCCGGACGGCATTTGGCATTGAATTTGCACGCTCAGAGGCTTTC---
       .. :.: :.:..:::  :::.:: .:::. :: ..: ::.... : :: ::.:::::.: :::
       CAGTCGGCGGGAGGGACGAGAAACAAAGACTGCCCAAGCAATAATGGTGGCCTGCACAGAAGGGTTCAGT
             730       740       750       760       770       780       790


           890       900       910       920       930       940       950
inputs CAGA---AGGGTGGAAGGAAAGGAGCCAAGAAGGTGATGATTGTCATCACAGATGGGGAGTCCCACGACA
       ::.. : ::  :.  .::.. .:  : :: :: : :.::.: :.::::::::.:.:::.:::::::.
       CAGTCCCATGGGGGCCGACCCGAGGCTGCCAGGCTACTGGTGGTTGTCACTGATGGAGAGTCCCATGATG
           800       810       820       830       840       850       860


             960       970       980       990      1000      1010      1020
inputs GCCCAGA----CCTGGAGAA-GGTGATCCAGCAAAGCGAAAGAGACAACGTAACAAGATATGCGGTGGCC
       :   ..::   :::: :: :  .... ::.: .:.:: ...:::  ::..::: : :::: ..: ::
       GAGAGGAGCTTCCTGCAGCACTAAAGGCCTGTGAGGC-TGGAAGA----GTGACACGCTATGGGATTGCA
           870       880       890       900       910       920       930


           1030      1040      1050      1060      1070      1080      1090
inputs GTCCTGGGCTACTACAACCGCAGGGGGATCAATCCAGAAACTTTTCTAAATGAAATCAAATACATCGCCA
       :::::  ::  :::::  .:::  :: .:  .::::  .. .:::: ::.:..:::: :.:.  :: ::::
       GTCCTTGGTCACTACCTCCGGCGGCAGCGAGATCCCAGCTCTTTCCTGAGAGAAATTAGAACTATTGCCA
           940       950       960       970       980       990      1000


           1100      1110      1120      1130      1140      1150      1160
inputs GTGACCCTGATGACAAGCACTTCTTCAATGTCACTGATGAGGCTGCCTTGAAGGACATTGTCGATGCCCT
       :::: ::.::::: ..  .::::::::::.::::::::::::::: ::: :::::::: ::::: :: :
       GTGATCCAGATGAGCGATTCTTCTTCAATGTCACAGATGAGGCTGCTCTGACTGACATTGTGGATGCACT
           1010      1020      1030      1040      1050      1060      1070


           1170      1180      1190      1200      1210      1220
inputs GGGGGACAGAATCTTCAGCCTGGAAGGCACC--AACA-AGAACGAGACCTCCTTTGGGCTGGAGATGTCA
       .::.:: :.:: :: .:::: :::::  .::  ..:: :.::::.: ::::::::::::::::.:::::::.
       AGGAGATCGGATTTTTGGCCTTGAAGGGTCCCATGCAGAAAACGAAAGCTCCTTTGGGCTGGAAATGTCT
           1080      1090      1100      1110      1120      1130      1140


       1230      1240      1250      1260      1270      1280      1290
inputs CAGACGGGCTTTTCCTCGCACGTGGTGGAGGATGGGGTTCTGCTGGGAGCCGTCGGTGCCTATGACTGGA
       ::::  :: :: :::.: ::  : :..::::::::::::.::: .:::: :: :: ::::::::::::::.
       CAGATTGGTTTCTCCACTCATCGGCTAAAGGATGGGATTCTTTTTGGGATGGTGGGGGCCTATGACTGGG
       1150      1160      1170      1180      1190      1200      1210


       1300      1310      1320      1330      1340      1350      1360
inputs ATGGAGCTGTGCTAAAG-GAGACGAGTGCCGGGAAGGTCATTCCTCTCCGCGAGTCCTACCTGAAA--GA
       ...: :::::::::...  .:: :.. ::: ::: : ::.:::: :: :::.. :: ::: ::.. ::
       GAGGCTCTGTGCTATGGCTTGAAGGAGGCC---ACCGCCTTTTCCCCCCACGAATGGCACTGGAAGACGA
       1220      1230      1240      1250      1260      1270
```

FIG. 3B

```
           1370      1380      1390      1400      1410       1420      1430
inputs GTTCCCCGAGGAGCTCAAGAACCATGGTGCATACCTGGGGTACACAGTCACATCGGTCGTG-TCCTCCAG
       ::::::::    :  .::  :::::::::: .:: :::::::::: :::.:.:: .:.:: :  .::  :.:
       GTTCCCCCCTGCACTGCAGAACCATGCAGCCTACCTGGGTTACTCTGTTTCTTC---CATGCTTTTGCGG
           1280      1290      1300      1310      1320      1330      1340


           1440      1450      1460      1470      1480      1490
inputs GCAGG--GGCGAGTGTACGTGGCCGGAGCCCCCCGGTTCAACCACACG-GGCAAGGTCATC-CTGTTCAC
       :  .:: :  ::   :::.  :  : ::.:: :: ::.:: :. :: .:: :: ::.:::::: :  : ::
       GGTGGACGCCGCCTGTTTCTCTCTGGGGCTCCTCGATTTAGACA-TCGAGGAAAAGTCATCGCCTTCCAG
           1350      1360      1370      1380      1390      1400      1410


           1500      1510      1520      1530      1540      1550      1560
inputs CATGCACAACAACCGGAGCC-TCACCATCCACCAGGCTATGCGGGGCCAGCAGATAGGCTCTTACTTTGG
       ::.: . ::  ::. ::.::   ::  .:   :::: ::::::::. :::: ::::..:: ::.:::::::
       CTTAAGAAAGAT--GGGGCTGTGAGGGTTGCCCAGAGCCTCCAGGGGGAGCAGATTGGTTCATACTTTGG
           1420      1430      1440      1450      1460      1470      1480


           1570      1580      1590      1600      1610      1620      1630
inputs GAGTGAAATCACCTCGGTGGACATCGACGGCGACGGCGTGACTGATGTCCTGCTGGTGGGCGCACCCATG
       :::::.  ::.  : :.  :::: ::  .:  :: :::  :.:::::::: :.:: :::. ::  ::::::
       CAGTGAGCTCTGCCCATTGGATACAGATAGGGATGGAACAACTGATGTCTTACTTGTGGCTGCCCCCATG
           1490      1500      1510      1520      1530      1540      1550


           1640      1650      1660      1670      1680      1690      1700
inputs TACTTCAACGAGGGCCGTGAGCGAGGCAAGGTGTACGTCTATGAGCTGAGACAGAA---CCGGTTTGTTT
       :.:  :.  :.::  ::  .::.:.::: ::.  .:  :::: :.:.: ::...:.:.   ::.:  . ::
       TTC--CT--GGGACCCCAGAACAAGG-AAACAGGACGTGTTTATG-TGTATCTGGTAGGCCAGCA-GTCC
               1560      1570      1580      1590      1600      1610


           1710      1720      1730      1740      1750      1760
inputs ATAACGGAACGCTAAAGGATTCA-CA--CAGTTACC----AGAATGCCCGATTTGGGTCCTCCATTGCCT
       .:: :. :. ::.:.: :: ::..: :: ::.:::. :: ::::.:: ::.:::: .:: :  :::: :
       TTGCTGACCCTCCAAGGAACACTTCAGCCAGAACCCCCCCAGGATGCTCGGTTTGGCTTTGCCATGGGAG
           1620      1630      1640      1650      1660      1670      1680


           1770      1780      1790      1800      1810      1820      1830
inputs CAGTTCGAGACCTCAACCAGGATTCCTACAATGACGTGGTGGTGGGAGCCCCCCTGGAGGACAACCACGC
       :. ::: .:: :: ::::::.:.  . ::: :::: :. :::::.: :: :::::. .. :::
       CTCTTCCTGATCTGAACCAAGATGGTTTTGCTGATGTGGCTGTGGGGGCGCCTCTGGAAGATGGGCACCA
           1690      1700      1710      1720      1730      1740      1750


           1840      1850      1860      1870      1880      1890      1900
inputs AGGAGCCATCTACATCTTCCACGGCTTCCGAGGCAGCATCCTGAAGACACCTAAGCAGAGAATCACAGCC
       .:::::   :  ::: :  :.:::: ::  .  ::...: .: .:: :     .:::. :::::.:: .:::
       GGGAGCACTGTACCTGTACCATGGAACCCAGAGTGGAGTCAGGCCCCATCCTGCCCAGAGGATTGCTGCT
           1760      1770      1780      1790      1800      1810      1820


           1910      1920      1930      1940      1950      1960      1970
inputs TCAGAGCTGGCTACCGGCCTCCAG-TATTTTGGCTGCAGCATCCACGGGCAATTGGACCTCAATGAGGAT
       :  . :. :: :: ::   :::  :::  :  :::::.: :  .: :  :: .::: ::: ::.:::::.:::
       GCC-TCCATGCCACATGCCCTCAGCTACTTTGGCCGAAGTGTGGATGGTCGGCTAGATCTGGATGGAGAT
           1830      1840      1850      1860      1870      1880      1890


           1980      1990      2000      2010      2020      2030      2040
inputs GGGCTCATCGACCTGGCAGTGGGAGCCCTTGGCAACGCTGTGATTCTGTGGTCCCGCCCAGTGGTTCAGA
       :.. ::  .::::: ::::::.:::::::.:::::.:: . . :: .:  :::::: . .: :: ::  ::
       GATCTGGTCGATGTGGCTGTGGGTGCCCAGGGGGGCAGCCATCCTGCTCAGCTCCCCGGCCCATTGTCCA--
           1900      1910      1920      1930      1940      1950      1960


           2050      2060      2070      2080      2090      2100      2110
inputs TCAATGCCAGCCTCCACTTTGAGCCATCCAAG--ATCAACATCTTCCACAGAGACTGCAAGCGCAGTGGC
       ::....  :: . ..: :..     ...:: ::: :.   :.::.:  :: .:::. ::..: :::  . :::
       TCTGACCCCATCACTGGAGGTGACCCCACAGGCCATCAGTGTGGTTCAGAGGGACTGTAGGCGGCGAGGC
           1970 .    1980      1990      2000      2010       2020      2030
```

FIG. 3C

```
            2120        2130        2140        2150        2160        2170        2180
inputs AGGGATGCCACCTGCCTGGCCGCCTTCCTCTGCTTCACGCCCATCTTCCTGGCACCCCATTTC-CAAACA
       ..::.::  .  :::  :::.:  ::    :::  :::::: :.    :: :: :.... :
       CAAGAAGCAGTCTGTCTGACTGCAGCCCTTTGCTTC-CAAGTGACCTCCCGT-ACTCCTGGTCGCTGGGA
            2040        2050        2060        2070        2080        2090

            2190        2200        2210        2220        2230        2240        2250
inputs ACAACTGTT-GGCATCAGATACAACGCCACCATGGATGAGAGG-CGGTATACAC-CGAG-GGCCCACCTG
       ..::  :..::  .:::  ::.:.:: :::  .:    ::::::::..:: : :  . .:: ..: .::  .
       TCACCAATTCTACATGAGGTTCACCGCATCACTGGATGAATGGACTGCTGGGGCACGTGCAGCATTTGAT
       2100.     2110        2120        2130        2140        2150        2160

            2260        2270        2280        2290        2300        2310
inputs GACGAGGGCGGGGACCGATTCACCAACAGAGCCGTACTGCTCTCCTCCGGCCAGGAGCTCT-GTGAGCGG
       :.:    :::  ..  :.:   ::  ..::: ..:.:.:  .  .   :.:: :::::::.:
       GGCTCTGGCCAGAG--GTTGTCCCCTCGGAGGC-TCCGGCTCAG-TGTGGGGAATGTCACTTGTGAGCAG
       2170        2180        2190        2200        2210        2220        2230

       2320        2330        2340        2350        2360        2370        2380
inputs ATCAACTTCCATGTCCTGGACACTGCTGACTACGTGAAGCCAGTGACCTTCTCAGTCGAGTATTCCCTGG
       :    :::::::::::  :::::  ::. :.:: :::  .  .::::::::.:::: .::.:  .  :.: : ::
       CTACACTTCCATGTGCTGGATACATCAGATTACCTCCGGCCAGTGGCCTTGACTGTGACCTTTGCCTTGG
       2240        2250        2260        2270        2280        2290        2300

       2390        2400        2410        2420        2430        2440        2450
inputs A--GGAC--CCTGACCATGGCCCCATGCTGGACGACGGCTGGCCCACCACTCT-CAGAGTCTCGGTGCCC
       :  .  .::  .  ..::::.:  ::  ::  ::::  .:::::::..:: : ::.:.  .:  :::  :::
       ACAATACTACAAAGCCAGGGCCT-GTGCTGAATGAGGGCTCACCCACCTCTATACAAAAGCT-GGTCCCC
       2310        2320        2330        2340·       2350        2360        2370

            2460        2470        2480        2490        2500        2510        2520
inputs TTCTGGAACGGCTGCAATGAGGATGAGCACTGTGTCCCTGACCTTGTGTTGGATGCCCGGAGTGACCTGC
       ::::         :::  :::::         ::::  :::::::  .:: .:  :.   :     .:::::
       TTCT---------~CAA--AGGAT------TGTGGCCCTGACAATGAAT--GTGTCAC----AGACCTG-
                    2380        2390        2400        2410

            2530        2540        2550        2560        2570        2580        2590
inputs CCACGGCCATGGAGTACTGCCAGAGGGTGCTGAGGAAGCCTGCGCAGGACTGCTCCGCATACACGCTGTC
                     ::.::  :  ::.:...     :..:::    :.  :.  :::      ::::          :  . .: :
       -------------GTGCTTCAAGTGAA---TATGGA---CATCAGAGG----CTCC------AGGAAGGC
                    2420        2430        2440        2450

            2600        2610        2620        2630        2640        2650        2660
inputs CTTCGACACCACAGTCTTCATCATAGAGAGCACACGCCAGCGAGTGGCGGTGGAGGCCACACTGGAGAAC
       :             :::  .  :  .:..:..: ::    :: :.: .:::.  ::::.
       C-------CCA-TTTGTGGTTCGAGGTG-GC---CGGCGGAAAGTGCTGGTATCTACAACTCTGGAGAAC
       2460        2470        2480        2490        2500        2510

            2670        2680        2690        2700        2710        2720        2730
inputs AGGGGCGAGAACGCCTACAGCACGGTCCTAAATATCTCGCAGTCAGCAAACCTGCAGTTTGCCAG-CTTG
       ::....  ::..::  :: :  :::::.  :::.  :::.::::.:::.    .  ::.. :::::: ::  :: : :::::  ::
       AGAAAGGAAAATGCTTACAATACGAGCCTGAGTATCATCTTCTCTAGAAACCTCCACCTGGCCAGTCTCA
       2520        2530        2540        2550        2560        2570        2580

            2740·       2750        2760        2770        2780        2790        2800·
inputs ATCC--AGAAGGAGGACTCAGACGGTAGCATTGAGTGTGT-GAACGAGGAGAGGAGGCTCCAGAAGCAAG
       :::  :::..:::  :...   :: :  :: :: ::..::: .    .:  :: :::.    ::
       CTCCTCAGAGAGAGAGCCCAAT--AAAG-GTGGAATGTGCCGCCCCTTCTGCTCATGC-CCG---GC---
       2590        2600        2610        2620        2630        2640

            2810        2820  ·      2830        2840        2850        2860        2870
inputs TCTGCAACGTCAGCTATCCCTTCTTCCGGGCCAAGGCCAAGGTGGCTTTCCGTCTTGATTTTGAGTTCAG
       :::::::.  ::  .:    ::::    :::::::..   ...::::::::::::  ::  :    :::::::::  ::
       TCTGCAGTGTGGGGCATCCTGTCTTCCAGACTGGAGCCAAGGTGACCTTTCTGCTAGAGTTTGAGTTTAG
       2650        2660        2670        2680        2690        2700        2710
```

FIG. 3D

```
          2880      2890      2900      2910      2920      2930
inputs CAAATCCATCT-TCCTACACCACCTGGAGATCGAGCTCGCTGCAGGCAGTGACAGTAATGAGCGGGACAG
       :. ::: ::: :::::  ..::  . .:::: .:::: ::::.::::::::::
       CTGCTCC-TCTCTCCTGAGCCAGGTCTTTGGGAAGCTGACTGCCAGCAGTGACAG--------------
          2720      2730      2740      2750      2760      2770

          2940      2950      2960      2970      2980      2990      3000
inputs CACCAAGGAAGACAACGTGGCCCCCTTACGCTTCCACCTCAAATACGAGGCTGACGTCCTCTTCACCAGG
         ::..:..:: :  :::: ::::: :   :. .::.:: ::. :: :::.
   \\  --CCTGGAGAGAAA---TGGCACCCTT---------CAAGAAAAC-------ACAGCC-------CAGA
               2780      2790              2800           2810

          3010      3020      3030      3040      3050      3060      3070
inputs AGCAGCAGCCTGAGCCACTACGAGGTCAAGCCCAACAGCTCGCTGGAGAGATACGATGGTATCGGGCCTC
       :  ::::::: ::..: :. . .:::::: ::  ::: ::: :.: :.:: :.: ::
       --CCTCAGCCT-----ACATCCAATATGAGCCCCAC--CTC-CTG------TTCTCTAGT---GAGTCTA
          2820           2830      2840         2850      2860

          3080      3090      3100      3110      3120      3130      3140
inputs CCTTCAGCTGCATCTTCAGGATCCAGAACTTGGGCTTGTTCCCCATCCACGGGATGATGATGAAGATCAC
       ::    :::::   ::           :::      :::         ::::.:.:::
       CC-----CTGCA----------CC---------GCT----------------------ATGAGGTTCAC
          2870                                              2880

          3150      3160      3170      3180      3190      3200      3210
inputs CATTCCCATCGCCACCAGGAGCGGCAACCGCCTACTGAAGCTGAGGGACTTCCTCACGGACGAGGCGAAC
       :          ::  .::..:     :: ::    :.:.:::  ::::       :::
       C-----------CATATGGGAC-----CCTCC--------CAGTGGG---TCCT---------GGC----
                   2890            2900              2910

          3220      3230      3240      3250      3260      3270      3280
inputs ACGTCCTGTAACATCTGGGGCAATAGCACTGAGTACCGGCCCACCCCAGTGGAGGAAGACTTGCGTCGTG
       ::.:.. .:.:... :  . :: . :    ::.:. .::: :::
       ----CCAGAAT--TCAAAACCACT--CTCAGGGTTCAGAACCT------------AGGCT-GCTATGTG
          2920      2930      2940      2950                      2960

          3290      3300      3310      3320      3330      3340      3350
inputs CTCCACAGCTGAATCACAGCAACTCTGATGTCGTCTCCATCAACTGCAATATACGGCTGGTCCCCAACCA
       ::  . .::    :..:: ::..::.  : :  :::::.::  :.:..:    :::          :::
       GTCAGTGGC-----CTCATCATCTCA---GCC--CTCCTTC--CAGCTGT----GGC----------CCA
          2970           2980         2990      3000

          3360      3370      3380      3390      3400      3410      3420
inputs GGAAATCAATTTCCATCTACTGGGGAACCTGTGGTTGAGGTCCCTAAAAGCACTCAAGTACAAATCCATG
       :.... :::::.:          ..::.:        : ::..     ::::::: ::. ..:
       TGGGGGCAATTAC-----------TTCCTAT---------CACTGT-----CTCAAGT-CATCAC----
          3010      3020                    TTCCTAT         3030      3040

          3430      3440      3450      3460      3470      3480      3490
inputs AAAATCATGGTCAACGCAGCCTTGCAGAGGCAGTTCCACAGCCCCTTCATCTTCCGTGAGGAGGATCCCA
       .:.::. : ::: ::.:: :.:        ::::.::.          ..::          ::.:T:~
       --TAACAATG-CAA-GCTGCATAG----TGCAGAACCT----------GACT-----------GAACCC-
          3050      3060         3070                  3080

          3500      3510      3520      3530      3540      3550      3560
inputs GCCGCCAGATCGTGTTTGAGATCTCCAAGCAAGAGGACTGGCAGGTCCCCATCTGGATCATTGTAGGCAG
                                    :::.:: :::         ::: :::          .:
       ------------------------CCAGGC-----------------CCACCTG-------------TG
                                3090                    3100

          3570      3580      3590      3600      3610      3620      3630
inputs CACCCTGGGGGGCCTCCTACTGCTGGCCCTGCTGGTCCTGGCACTGTGGAAGCTCGGCTTCTTTAGAAGT
       :: ::.:.::.::. ::        ..: :. :.... :.:.:. ..::::.:::. .:: :    :::
       CATCCAGAGGAGCTTC--------AACACA-CAAA--CAGACTGAATGGGAGCAATACT-C------AGT
          3110 ·              3120        3130      3140      3150
```

FIG. 3E

```
        3640      3650      3660      3670      3680      3690      3700
inputs GCCAGGCGCAGGAGGGAGCCTGGTCTGGACCCCACCCCCAAAGTGCTGGAGTGAGGCTCCAGAGGAGACT
       : :::: :   :  ..::.:::.   ::  :. :       .:::::            ::.::.:.:::
       GTCAGGTG--GTGAGGTGCCA--CCTTGGGC-----------AGCTGG----------CAAAGGGGACT
          3160      3170      3180                  3190

        3710      3720      3730      3740      3750      3760      3770
inputs TTGAGTTGATGGGGGCCAGGACACCAGTCCAGGTAGTGTTGAGACCCAGGCCTGTGGCCCCACCGAGCTG
                                  .::::...  ::::::.      ::. :.      ::
       G------------------AGGTCTC-----------TGTTGGA-------CTATTG-------AGGCTG
                            3200              3210           3220

        3780      3790      3800      3810      3820      3830      3840
inputs GAGCGGAGAGGAAGCCAGCTGGCTTTGCACTTGACCTCATCTCCCGAGCAATGGCGCCTGCTCCCTCCAG
       :. :. :..:.:.   . :   ..:: :..:. :::. :::: ..:..:::  :.:.::.:::: :  ::
       GTTCACAATGAATTTTTCC--GAAGAGC-CAAGT--TCAAGTCCCTGACGGTGGT--CAGCACCTTTGAG
          3230      3240       3250        3260      3270      3280

        3850      3860      3870      3880      3890      3900      3910
inputs AATGGAACTCAAGCTGGTTTTAAGTGGAACTGCCCTACTGGGAGACTGGGACACCTTTAACACAGACCCC
       .  .:::::  :.:         :.. :: : :: :::::.:       :::.       :: ..: :.      :::
       CTGGGAAC-CGA---------AGAGGGCAGTGTCCTACAG-----CTGA-----CTGAAGC-CT---CCC
          3290              3300      3310           3320         3330

        3920      3930      3940      3950      3960      3970      3980
inputs TAGGGATTTAAAGGGACACCCCTACACACACCCAGGCCCACGCCAAGGCCTCCCTCAGGCTCTGTGGAGG
       .  ::: :.::.                                      ::::.:. :.::----GAGGTGG
       GTTGGAGTGAGAG----------------------------------------CCTCTTG----GAGGTGG
          3340                                              3350      3360

        3990      4000      4010      4020      4030      4040      4050
inputs GCATTTGCTGCCCCAGCTACTAAGGTGCTAGGAATTCGTAATCATCCCCATCCTCCAGAGAAACCCAGGG
       : :.: :::.::           :::                :::  ::::  ::.:..:.:.::
       -----TTCAGACCCGGC----------CTA-------------TCCTCATCTCCCTGTGGATCC-----
            3370                                   3380      3390

        4060      4070      4080      4090      4100      4110      4120
inputs AGGAAGACTGTAAATACGAACCCAATCTGCACACTCCAGGCCTCTAGTTCCAGAAGGATCCAAGACAAAA
       .  ..::.:.::.                          :::.::            ::.::.: :
       TCATAGGCAGTG----------------------TCCTGG------------GAGGGTTGC---------
          3400                            3410              3420

        4130      4140      4150      4160      4170      4180      4190
inputs CAGATCTGAATTCTGCCCCTTTTCTCTCACCCATCCCCACCCCTCCATTGGCTCCCAAGTCACACCCACTCC
       ::       ::::    :: :::::  :.:  :.:   ::. ::: : ...      :: :
       ----TC------CTGC---TTGCTCTC---CTTGTCTTCT-GCCTGTGGAAGCTTGG---------CTTC
          3430          3440        3450      3460

        4200      4210      4220      4230      4240      4250      4260
inputs CTTCCCCATAGATAGGCCCCTGGGGCTCCTGAAGAATGAACCCAAGAGCAAGGGCTTGATGGTGACAGCT
       :: :::::::....:.. :::::.::    :::::..:     ::::  :::.:  ..:.:::.: ::. :
       TTTGCCCATAAGAAAATCCCTGAGG------AAGAAAAA-----AGAG-AAGAG-AAGTTGGAG-CAA-T
          3470      3480      3490           3500           3510      3520

        4270      4280      4290      4300      4310      4320      4330
inputs GCAAGCCAGGGATGAAGAAAGACTCTGAGATGTGGAGACTGATGGCCAGGCAAGTGGGACCAGGATACTG
       :           :. :..::: ::..:. :::......::     ::      ::.::::.      :. :..::
       G------AATGTAGAATAAGGG-TCTAGAAAGT----CCTC----CCTGGCAG-------CTTTCTTCAA
             3530      3540           3550      3560               3570

        4340      4350      4360      4370      4380      4390      4400
inputs GACGCTGTCCTGAGATGAGAGGTAGCCGGGCTCTGCACCCACGTGCATTCACATTGACCGCAACTCACAC
       ::  .:: :..::..: :::::::. ::::::.:       .:: ..: :: ...:::: .:: . ::
       GAGACTTGCATAAAAGCAGAGGTTTGGGGGCTCAG-------ATG-GGACA-AGAAGCCGCCTCTGG-AC
             3580      3590      3600           3610      3620      3630
```

FIG. 3F

```
          4410      4420      4430      4440      4450      4460      4470
inputs ATTCCCCCACCAGCTGCAGCCCCCTTGCTCTCAGCTGCCAACCCTCCCGGGTCACTTTTGTTCCCAGGTA(
       ..::  :::           :::  ::       :::.:::.. : :      : :::::::::.         ::
       TATCTCCC--------CAGACC---------AGCAGCCTGACTT-----G--ACTTTTGA------GT-(
                  3640                3650              3660            367:

          4480      4490      4500      4510      4520      4530      4540
inputs CTCATGGGAAGCATGTGGATGACACAATCCCTGGGGCTGTG-CATTCCCACGTCTTCTTGCTGCAGCCT.
        ::   .::::.::.  :::              ::.:.: ::.:  :.:: ::.:.      .:..: .: :::
         CT--AGGGATGCTGCTGG-----------CTAGAGATGAGGCTTTACCTCAG--ACAAGAAG-AGCTG
                  3680                    3690      3700      3710      3720

          4550      4560      4570      4580      4590      4600      4610
inputs CCCCTAGACATGGACGCACTGGCCTGGCTGCAGCTGGGCAGCAGGGGTAGGGGTAGGGAGCCTCCCCTC
       : ::.:.::           :.:::. :::                                       :::.:
       CACCAAAAC---------TAGCCATGCT----------------------------------CCCAC
             3730               3740

          4620      4630      4640      4650      4660      4670      4680
inputs CTGTATCACCCCCTCCCTACACACACACACACACACACACACACTGCCTCCCATCCTTCCCTCAT
        ::                                                  :::: :::: :::: :::::.:
       CT-----------------------------------------------CTGCTTCCC-TCCT--CCTCGT
       3750                                                       3760.       377

          4690      4700      4710      4720      4730      4740      4750
inputs CCCGCCAGTGCACAGGGAAGGGCTTGGCCAGCGCTGTTGAGGGGTCCCCTCTGGAATGCACTGAATAAA
       : :  .:: :            :.:.:::::.:.::    :::     :: : :..::
       ATC-CTGGTTC-----------CATAGCCAACACTG----GGG-----CTTTTGTTTG-----------
              3780                    3790              3800

          4760      4770      4780      4790      4800      4810      4820
inputs CACGTGCAAGGACTCCCGGAGCCTGTGCAGCCTTGGTGGCAAATATCTCATCTGCCGGCCCCCAGGACA
                        ::. ::: :                    ::::            :::::::
       ----------------GGGTCCTTT----------------TATC-------------CCCAGGA--
                       3810                     3820

          4830      4840      4850      4860      4870      4880      4890
inputs GTGGTATGACCAGTGATAATGCCCCAAGGACAAGGGGCGTGCCTGGCGCCCAGTGGAGTAATTTATGCC
             : ::.:.::.                                              :::.::::
       -------ATCAATAATT----------------------------------------------TTTTTGCC
              3830                                                      3840

          4900      4910      4920      4930      4940      4950      4960
inputs TAGTCTTGTTTTGAGGTAGAAATGCAAGGGGGACACATGAAAGGCATCAGTCCCCCTGTGCATAGTACC
               :::.:.:::                         :::.::   : ::       ::. :.:.::
       -----------AGGAAAAAA----------------AAAAGC----GGCC------GCG-AATTCC
                  3850                           3860           3870

          4970      4980      4990      5000      5010      5020      5030
inputs CCTTTACTGTCGTATTTTTGAAAAATTAAAAATACAGTGTTTAAAAACAAAAAAAAAAAAAAAAGGGCC
            :.::                                         ::.:.
       -------TATC---------------------------AAGCT--------------------
                                                    3880

          5040
inputs CCGC

       ----
```

FIG. 3G

```
            10        20        30        40        50        60        7
inputs  MDLPRGLVVAWALSLWPGFTDTFNMDTRKPRVIPGSRTAFFGYTVQQHDISGNKWLVVGAPLETNGYQK
        :.::     .  .: .  :.  ::.: .  ::..::. :  :::.:  ::  .:...:.::: .  ..
        MELPFVTHLFLPLVFLTGLCSPFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPWDGPSGDK
            10        20        30        40        50        60        7

            80        90       100       110       120       130
inputs  GDVYKCPV--IHGN-CTKLNLGRVTLSNVSERKDNMRLGLSLATNPKDNSFLACSPLWSHECGSSYYT
        :::::::  ::..: ..:.:  :.: . ::..::. ... ::....:::: ...
        GDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHPAVNMHLGMSLLETDGDGGFMACAPLWSRACGSSVFS
            80        90       100       110       120       130       14

           140       150       160       170       180       190       200
inputs  MCSRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGSNSIYPWVEVQHFLINILKKFYIGPGQIQVGVVQ
        .:..::.     ..:::. :::.::::.::::::::::: :::  :: .. :..:..:::::.:::
        ICARVDASFQPQGSLAPTAQRCPTYMDVVIVLDGSNSIYPWSEVQTFLRRLVGKLFIDPEQIQVGLVQ
           150       160       170       180       190       200       2:

           210       220       230       240       250       260       270
inputs  EDVVHEFHLNDYRSVKDVVEAASHIEQRGGTETRTAFGIEFARSEAFQK--GGRKGAKKVMIVITDGE
        :. :::  :.:.:.  .:: ::  .. .:  ::.:: .:  :  .:.:  .  ::: : ....:.::::
        ESPVHEWSLGDFRTKEEVVRAAKNLSRREGRETKTAQAIMVACTEGFSQSHGGRPEAARLLVVVTDGE
           220       230       240       250       260       270       2:

           280       290       300       310       320       330       340
inputs  DSPDLEKVIQQSERDNVTRYAVAVLGYYNRRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIVD
        :. .:  .: . :::::: ::.:::: :.:.. .: .. :..: ::::::::::: ::::::::: :::
        DGEELPAALKACEAGRVTRYGIAVLGHYLRRQRDPSSFLREIRTIASDPDERFFFNVTDEAALTDIVD
           290       300       310       320       330       340       3!

           350       360       370       380       390       400       410
inputs  GDRIFSLEGTN-KNETSFGLEMSQTGFSSHVVEDGVLLGAVGAYDWNGAVLKETSAGKVIPLRESYLK
        :::::.:::.. :..:.::::::::.:::.:  . ::.:.: :::::::::  ...: : .
        GDRIFGLEGSHAENESSFGLEMSQIGFSTHRLKDGILFGMVGAYDWGGSVLWLEGGHRLFPPRMALED
           360       370       380       390       400       410       4:

           420       430       440       450       460       470       480
inputs  PEELKNHGAYLGYTVTSVVSSRQGRVYVAGAPRFNHTGKVILFTMHNNRSLTIHQAMRGQQIGSYFGS
        : .:.::.:::::.:.:::.:.:. :...:::: :: ::::. :::: : ..:..:.:::::::::::
        PPALQNHAAYLGYSVSSMLLRGGRRLFLSGAPRFRHRGKVIAFQLKKDGAVRVAQSLQGEQIGSYFGS
           430       440       450       460       470       480       4.

           490       500       510       520       530       540       550
inputs  TSVDIDGDGVTDVLLVGAPMYFN-EGRERGKVYVYEL-RQNRFVYNGTLKDSHSYQNARFGSSIASVR
        ..:.: ::.:::::::::.:.::: ....:::: :..::.:: ...::: . :::..: :.::.:.:.
        CPLDTDRDGTTDVLLVAAPMFLGPQNKETGRVYVYLVGQQSLLTLQGTLQPEPP-QDARFGFAMGALP
           500       510       520       530       540       550

           560       570       580       590       600       610       620
inputs  NQDSYNDVVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRITASELATGLQYFGCSIHGQLDLNEDGLI
        :::...:.:.::::::.: ::..::.::: .... :  ::::.:.. ..: ::: :..:.:: :...
        NQDGFADVAVGAPLEDGHQGALYLYHGTQSGVRPHPAQRIAAASMPHALSYFGRSVDGRLDLDGDDLV
        560       570       580       590       600       610       620

           630       640       650       660       670       680       690
inputs  AVGALGNAVILWSRPVVQINASLHFEPSKINIFHRDCKRSGRDATCLAAFLCFTPIFLAPHFQTTTVG
        :::: :..:.: ::::.:..::: :   ::.  .:::.: :..:.::.: :::  . .:
        AVGAQGAAILLSSRPIVHLTPSLEVTPQAISVVQRDCRRRGQEAVCLTAALCFQVTSRTPGRWDHQFY
        630       640       650       660       670       680       690
```

FIG. 4A

```
        700        710        720        730        740        750        7(
inputs YNATMDERRYTPRAHLDEGGDRFTNRAVLLSSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPD-I
       ..:...::.   .:: .:..:....  :  . :: :.  ::....::::::.::..::...: ...:..
       FTASLDEWTAGARAAFDGSGQRLSPRRLRLSVGNVTCEQLHFHVLDTSDYLRPVALTVTFALDNTTKI
        700        710        720        730        740        750        760


        770        780        790        800        810        820        8:
inputs MLDDGWPTTLRVSVPFWNGCNEDEHCVPDLVLDARSDLPTAMEYCQRVLRKPAQDCSAYTLSFDTTV:
        ..:...::..   ::: ...:. :. :: :::::.   :.           :  : ::..
       VLNEGSPTSIQKLVPFSKDCGPDNECVTDLVLQVNMDI-------RGSRKAP-------------I
        770        780        790        800        810


        840        850        860        870        880        890        9(
inputs ESTRQRVAVEATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRLQKQVCNVSYPI
        . :...: :  .:::::: :::::.: :.:  : ::..::...  .:  .::..   . ...::. :
       RGGRRKVLVSTTLENRKENAYNTSLSIIFSRNLHLASLTPQRESPIKVECAAPSA--HARLCSVGHP'
        820        830        840        850        860        870        880


        910        920        930        940        950        960        9'
inputs AKAKVAFRLDFEFSKSIFLHHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHYI
        . ::::.: :..:::: :  .: ..   :.:.:::  ::..: ...:.     ..:::...::. .:: .:.
       TGAKVTFLLEFEFSCSSLLSQVFGKLTASSDSLERNGTLQENTAQTSAYIQYEPHLLFSSESTLHRYI
        890        900        910        920        930        940        950


        980        990        1000       1010       1020       1030       10·
inputs PNSSLERYDGIGPPFSCIFRIQNLGLFPIHGMMMKITIPIATRSGNRLLKLRDFLTDEANTSCNIWGI
        : ..::    : :: :  ..:.:::: : :...  .  ::..::..  :..:.:.
       PYGTLPV--GPGPEFKTTLRVQNLGCYVVSGLIISALLPAVAHGGNYFLSLSQVITN--NASCIV-QI
        960        970        980        990        1000       1010       1(


        1050       1060       1070       1080       1090       1100
inputs EYRPTPVE-EDLRRAPQLNHSNSDVVSINCNI-RLVPNQEINFHLLGNLWLRSLKALKYKSMKIMVN:
        :        ::  :..:.. .::  ::.. .  :...  .: ..  ::  .    .. :.::. .
       EPPGPPVHPEELQHTNRLNGSNTQCQVVRCHLGQLAKGTEVSVGLLRLVHNEFFRRAKFKSLTVVST:
        1030       1040       1050       1060       1070       1080       1(


 1110       1120       1130       1140       1150       1160       1170
inputs QRQFHSPFIFREEDPSRQIVFEISKQEDWQVPIWIIVGSTLGGLLLLALLVLALWKLGFFRSARRRRI.
       .   :  : .:..    .  .::. .    ...::...::.:::::::::::::. ::::::: :....
       GTEEGSVLQLTEASRWSESLLEVVQTRPILISLWILIGSVLGGLLLLALLVFCLWKLGFF--AHKKII
        1100       1110       1120       1130       1140       1150


 1180
inputs LDPTPKVLE
       :. .
       EKREEKLEQ
 1160
```

FIG. 4B

```
                                       M   D   F   P   R   G   L   L   V   A   W   T   L
GCGACACCCGCCGGCGGCAGCGGGCC    ATG GAC TTT CCC AGG GGC CTC CTG GTG GCC TGG ACC CTC      13
                                                                                        66

 S   L   W   P   G   F   T   D   T   F   N   M   D   T   F   R   N   P   V   I
AGC CTT TGG CCA GGA TTC ACG GAC ACC TTC AAT ATG GAT ACC TTC CGG AAT CCC GTC ATT        33
                                                                                       126

 A   G   P   S   A   A   F   F   G   Y   T   Q   V   Q   H   D   I   S   G   K
GCT GGT CCC AGT GCT GCC TTC TTT GGC TAC ACA CAG GTA CAG CAT GAT ATC AGT GGC AAG        53
                                                                                       186

 K   W   L   V   V   G   A   P   M   E   T   G   N   H   Q   K   T   G   D   V
AAG TGG TTG GTT GTT GGT GCC CCA ATG GAA ACC GGC AAT CAT CAG AAG ACA GGA GAC GTA        73
                                                                                       246

 Y   K   C   P   V   T   Q   G   N   C   T   K   L   N   L   G   R   V   T   L
TAC AAA TGT CCG GTA ACC CAG GGC AAC TGC ACC AAG CTC AAC CTG GGC AGG GTC ACT CTG        93
                                                                                       306

 S   N   V   S   E   R   K   D   N   M   R   L   G   L   S   L   A   T   N   P
TCC AAT GTG AGC GAG TCT AAG GAC AAC ATG CGC CTG GGC CTC AGC CTG GCC ACC AAC CCC       113
                                                                                       366

 K   D   N   S   F   L   A   C   S   P   L   S   L   W   H   E   C   G   S   Y
AAG GAC AAC AGC TTC CTG GCC TGC AGC CCT CTC TCG CTC TGG CAC GAG TGT GGA AGC TAC       133
                                                                                       426

 Y   T   T   G   M   C   S   R   V   N   S   W   N   F   R   S   K   T   V   A
TAC ACT ACT GGC ATG TGC TCA CGG GTC AAC TCC TGG AAC TTC AGA TCT AAG ACG GTA GCT       153
                                                                                       486

 P   A   L   Q   R   C   Q   T   Y   M   D   I   V   L   I   V   D   G   S   N
CCG GCA CTT CAG AGG TGC CAG ACT TAC ATG GAC ATT GTC TTA ATT GTC GAT GGC TCC AAC       173
                                                                                       546

 S   I   Y   P   W   V   E   V   Q   F   H   L   I   N   I   L   K   K   F   Y
AGC ATC TAC CCC TGG GTG GAG GTC CAA TTC CAC CTC ATC AAT ATC CTC AAA AAG TTC TAC       193
                                                                                       606

 I   G   P   G   Q   I   Q   V   G   I   V   Q   Y   G   E   D   A   V   H   E
ATT GGC CCC GGC CAG ATC CAG GTC GGA ATA GTC CAG TAT GGA GAA GAT GCC GTC CAT GAG       213
                                                                                       666
```

FIG. 5A

```
F   H   L   N   D   Y   R   S   V   K   D   V   E   A   S   H   I   E          233
TTC CAC CTT AAT GAC TAC AGG TCT GTA AAA GAT GTG GAA GCC AGC CAC ATT GAG        726

Q   R   G   G   T   E   T   R   A   F   I   E   E   F   A   R   S   E          253
CAG AGA GGA ACA GAG ACC CGC GCA TTT ATT GAA GAA TTT GCA CGC TCG GAG GCT        786

F   K   G   G   R   K   A   K   G   V   M   I   V   T   D   G   E              273
TTC AAG GGT GGA AGA AAA GCC AAA GGG GTG ATG ATT GTC ACG GAC GGG GAA            846

S   D   S   P   D   L   E   K   Q   S   E   K   D   N   V   T                  293
TCC GAC AGC CCA GAC CTG GAG AAG CAG AGC GAG AAG GAC AAC GTG ACC                906

R   Y   A   V   L   G   Y   N   R   R   G   I   N   P   E   T   F              313
AGA TAC GCT GTG TTG GGC TAC AAC CGC AGG GGG ATC AAT CCA GAG ACT TTT            966

L   N   E   I   K   Y   I   A   S   D   P   D   K   H   F   F   N   V   T      333
CTA AAT GAA ATC AAA TAC ATC GCC AGC GAC CCT GAC AAG CAC TTC TTC AAC GTC ACA   1026

D   E   A   A   L   K   D   V   D   A   L   G   D   I   F   S   L   E          353
GAT GAG GCG GCC CTG AAG GAC GTT GAC GCC CTT GGG GAC ATT TTC AGC TTG GAA        1086

G   T   N   K   N   E   T   S   F   G   L   E   M   S   Q   T   G   F   S   S  373
GGC ACC AAC AAG AAT GAG ACC TCT TTT GGG CTG GAG ATG TCG CAG ACT GGC TTT TCT TCA 1146
```

FIG. 5B

```
H   V   E   D   G   I   L   L   G   A   V   G   A   Y   D   W   N   G   A        393
CAT GTA GAG GAT GGG ATC CTG CTG GGA GCT GTG GGA GCC TAT GAC TGG AAC GGG GCG     1206

V   L   K   E   T   S   A   G   K   V   I   P   H   R   E   S   L   K   E        413
GTG CTG AAG GAG ACA AGT GCT GGC AAG GTG ATT CCT CAC CGA GAG TCC CTT AAG GAG     1266

F   P   E   E   L   K   N   H   A   A   Y   L   G   Y   T   V   S   V            433
TTT CCA GAG GAG CTG AAG AAC CAT GCA GCA TAC CTA GGG TAT ACG GTG TCG GTG         1326

S   R   Q   G   G   R   V   Y   Y   V   A   G   P   A   R   F   N   T   G        453
TCC AGG CAG GGG GGG CGG GTG TAT TAT GTG GCT GGA CCC GCC AGA TTC AAC ACT GGC     1386

V   L   F   S   M   N   H   N   R   S   L   T   I   H   Q   A   L   R            473
GTC CTA TTC AGC ATG AAC CAC AAC CGG AGC CTC ACC ATC CAC CAG GCT CTT CGG         1446

E   Q   I   G   S   Y   F   G   S   E   I   T   S   V   D   V   N   D   D   R    493
GAG CAG ATA GGC TCC TAC TTT GGG AGT GAG ATT ACC TCA GTG GAC GTC AAT GAT GAC AGA 1506

V   D   V   L   L   V   G   A   P   M   Y   F   S   E   R   E   G                513
GTG GAT GTG CTG CTG GTG GGT GCA CCC ATG TAC TTC AGC GAG CGA GAG GGC             1566

K   V   Y   N   Y   R   Q   N   R   F   V   Y   N   G   T   L   K   D            533
AAG GTG TAT AAC TAC CGG CAG AAC AGA TTT GTT TAT AAT GGC ACT CTG AAG GAT         1626

S   H   Y   Q   N   A   R   F   G   S   C   I   A   S   V   Q   D   L   N        553
TCC CAC TAC CAG AAC GCC CGG TTC GGG TCC TGC ATC GCC TCT GTT CAA GAC CTC AAC     1686

Q   D   S   Y   N   D   V   V   G   V   G   L   E   D   S   H   R   G   A        573
CAA GAT TCC TAC AAT GAT GTG GTG GGG GTG GGG CTG GAG GAC AGC CAC AGA GGG GCC     1746

I   Y   I   F   H   G   F   Q   T   N   L   K   P   K   M   Q   R   I   T        593
ATC TAC ATC TTC CAT GGC TTC CAA ACC AAT CTG AAG CCC AAG ATG CAG AGA ATA ACA     1806

A   S   E   L   A   P   G   L   Q   H   F   G   C   S   I   H   G   Q   L   D    613
GCC TCG GAG CTG GCT CCT GGC CTG CAA CAT TTT GGC TGC AGC ATC CAC GGA CAA CTG GAC 1866
```

FIG. 5C

```
    L   N   E   D   G   L   V   D   L   A   V   G   A   L   G   N   A   V   V   L     633
    CTC AAT GAG GAT GGG CTT GTG GAC CTA GCA GTG GGC GCC CTG GGC AAC GCT GTG GTT TTG   1926

    W   A   R   P   V   V   Q   I   N   A   S   L   H   F   E   P   S   K   I   N     653
    TGG GCC CGC CCT GTA GTT CAA ATC AAC GCC AGC CTC CAC TTT GAG CCT TCC AAG ATC AAC   1986

    I   F   H   K   D   C   K   R   N   G   R   D   A   T   C   L   A   F   L         673
    ATC TTC CAC AAG GAC TGC AAG CGC AAT GGC AGG GAT GCC ACC TGC CTG GCC TTC CTC       2046

    C   F   I   P   F   L   A   P   H   F   Q   T   A   V   G   I   R   Y             693
    TGC TTC ATA CCC ATC TTC CTG GCA CCC CAC CAA ACA GCA ACC GTC GGC ATC AGG TAC       2106

    N   A   T   M   D   E   R   Y   M   P   R   A   H   L   D   E   G   G   D         713
    AAT GCA ACC ATG GAT GAG AGG CGG TAT ATG CCG CGG GCA CAT CTG GAC GAG GGC GGA GAC   2166

    Q   F   T   N   R   A   V   L   S   G   S   Q   E   H   C   Q   R   I   N         733
    CAG TTC ACC AAC AGA GCC GTC CTC TCC GGC TCT CAG GAA CAC TGT CAA AGG ATC AAC       2226

    F   H   V   L   D   T   A   D   Y   V   K   P   V   A   F   S   V   E   Y   S     753
    TTC CAT GTC CTG GAC ACT GCA GAC TAC GTG AAG CCA GTG GCC TTC TCC GTG GAG TAC TCC   2286

    L   E   D   P   D   N   G   P   M   L   D   N   G   A   L   T   T   L   R   V     773
    CTA GAG GAC CCT GAC AAC GGC CCC ATG CTG GAC AAC GGC GCT CTG ACG CCC ACT CTC AGA GTG  2346
```

FIG. 5D

S V P F W N G C N E D H C V P D L V L
TCG GTG CCC TTC TGG AAT GGC TGT AAT GAG GAT CAC TGT GTC CCT GAC CTT GTA CTG
793
2406

D A R S D L P T A M E Y C Q R V L G R P
GAT GCT CGG AGT GAC CTG CCC ACT GCC ATG GAG TAC TGC CAG CGG GTG CTG GGG AGG CCG
813
2466

A Q D C S Y L S T D F T V I F I E I S
GCG CAG GAC TGC TCC TAC CTG AGC ACC GAC TTC ACT GTC ATC TTC ATA GAG ATC AGC
833
2526

T R R V A V E N R G A S R L A R L Q K D
ACG CGC CGG GTG GCA GTG GAG AAC AGA GGA GCC AGC AGA CTG GCT CGT CTG CAG AAG GAT
853
2586

A V L N I S Q S E N L Q A S L I Q K D S
GCC GTC CTC AAC ATC TCC CAG TCA GAG AAC CTT CAG GCC AGC CTT ATC CAG AAG GAT AGT
873
2646

D S D N S I C V E N E R A K A H R K V C
GAC TCA GAC AAC AGC ATC TGT GTG GAG AAC GAG AGG GCC AAG GCC CAC CGG AAA GTC TGC
893
2706

N S Y P Y P F F R A K L G A R L A K V E
AAC AGC TAC CCC TAC CCC TTC TTC AGA GCC AAG CTG GGT GCT CGT CTT GCC AAG GTC GAG
913
2766

F S K V F L L H I Q H L G A A G D F E S
TTC TCT AAG GTG TTT CTG CTG CAT ATC CAG CAC CTT GGT GCC GCC GGC GAC TTC GAG AGT
933
2826

H Q D S T A T L S S L R F H L K Y
CAC CAA GAC AGC ACC GCA ACC CTC AGC AGC CTG CGC TTC CAC CTC AAA TAT
953
2886

E A D V L F T T L F S L S H F E K A N
GAA GCA GAC GTC CTC TTT ACC ACC CTC TTT AGC CTG AGC CAC TTC GAG AAG GCA AAC
973
2946

S S E S Y D G I G P P F N C V F K V Q
AGC TCA GAG AGC TAT GAT GGC ATC GGA CCT CCC TTC AAC TGT GTT TTC AAG GTG CAG
993
3006

N L G F F P I H G V M K I T V P I A T
AAC CTG GGC TTT TTC CCC ATC CAC GGG GTG ATG AAG ATC ACT GTG CCC ATC GCC ACC
1013
3066

FIG. 5E

```
 R   G   G   N   R   L   L   M   L   R   D   F   F   T   D   Q   G   N   T   S      1033
AGG GGT GGC AAC CGC CTG CTG ATG TTG AGG GAC TTC TTC ACC GAC CAG GGC AAC ACA TCC     3126

 C   N   I   W   G   N   S   T   E   Y   R   S   T   P   T   E   E   D   L   S      1053
TGT AAC ATC TGG GGG AAC AGC ACA GAG TAC AGG AGT ACC CCA ACT GAA GAG GAT CTG AGC     3186

 H   A   P   Q   R   N   H   N   S   D   S   V   V   S   I   I   C   N   L   R      1073
CAT GCC CCA CAG AGG AAT CAC AAC AGC GAC TCC GTG GTC TCC ATC ATC TGC AAC CTG AGG     3246

 L   A   P   S   Q   E   T   S   F   Y   L   V   G   L   W   L   T   S   L          1093
CTG GCC CCC AGC CAG GAG ACC AGC TTC TAC CTG GTG GGA CTG TGG CTG ACA TCT CTG         3306

 K   A   L   K   Y   R   S   L   K   I   T   V   N   A   L   Q   R   Q   F          1113
AAA GCA CTC AAG TAC AGG TCT CTG AAG ATC ACA GTC AAC GCC CTG CAG AGG CAG TTC         3366

 H   S   P   F   I   F   R   E   E   D   P   S   R   Q   V   T   F   E   I   S      1133
CAC AGC CCC TTC ATC TTC CGA GAG GAA GAC CCT AGC CGC CAG GTC ACA TTT GAG ATC TCC     3426

 K   Q   D   W   Q   V   P   I   W   L   V   G   S   T   L   G   G   L              1153
AAG CAA GAC TGG CAG GTG CCC ATC TGG CTG GTG GGC AGC ACT CTG GGG GGC CTC             3486

 L   L   A   L   V   L   A   L   W   K   L   G   F   F   K   S   A   K              1173
TTG CTG GCA CTG GTC CTG GCC CTG TGG AAG CTC GGT TTT TTT AAA AGT GCC AAG             3546

 R   K   R   E   P   G   L   G   P   I   P   K   E   L   K   *                      1189
CGC AAG AGG GAA CCT GGC CTA GGC CCC ATC CCC AAA GAG CTC AAG TGA                     3594
```

FIG. 5F

```
GGATCCAGAGGAAGCTTCAAGTTGTTGGGGGCCCGACACCAGTCCGGGGGTGTGCAGGCCCGGGCTATGGCCTCAGAGC    3673

TGGAGCAAAGAGGACACCCATGGACTTTACATTGAGTTCATCTCTGGAGCAATGGCAACTGATCGCTGTGGGATGGACC    3752

TCCAGCCCGTGTCGGAGTGACTTCTCTACTAGAGGCCAAGACCTCCAGTACAGATCCCTAAGGACTAAAAGGGACCCT    3831

CCAAGACCCACCGGTGGCCTTCCCCCAAGGCTCTGTGGAGGAATCTGCCCCAAACACTAAGGTGCTAGGGTTTTG    3910

TTGTCTTGCCCACCTCTGAAGAAACTCAAGGGAAGCTAGCAAATACGAGCCCACTGTGCACGCCCTATCTGGACCTAT    3989

AGCTTCAGGACTCCATGGCCCGGGAGCATAACCTCCGCGCACCCTTACCCTTGCTCTGGTTCCCAAACCACAGCCA    4068

TCCCTTCTCCACAGATAGACCCCAAAGATCTCTATGAGTGTGACCCAGAGTTGTGTGCTTGATATGCCAGCTGCAGGCCA    4147

GTGATGGAAACACATCCAGGGACTGTGCCAAATGAACTGGGAGGAATCAGGATGCTCGGGTACCTTCCAAACATCAGAAA    4226

GCAGCAGGGCTCTCCAGCAGTAACCACCACCCACTCAATCCTGCTCTCCCCCGCAGCTCCTCTCAGCCGCTT    4305

TGGTTGCTAGGCAACTCGTGCTAGGGAGCCTGATAGGGAGCCTCTGTTTCTTACATCAGCACCTACACACACACAC    4384

ACTGCAGCTTGTCCCTGGTAGATGCACTGTGAGATGCACTGTGGGGAGAAGGGCTCCTGTA    4463

ACACACACACACACACACACACCGCGCCTTCTAGGCATGCACTGAACAAAGCCCATGCAAGGCCCTGGAACCTTAAACCTCAGAGG    4542

GGGGGCGCTGTGGGGGCCTTCTAGGCATGCACTGAACAAAGCCCATGCAAGGCCCTGGAACCTTAAACCTCAGAGG    4621

CAGATAGGCATCTCGCCGGCCAGCTCTGGGATAAGCCCAAGGACAAGGGCATGCCGGATGCCCAGCAGAGTAATTTA    4700

TGCCTTAACCTTTTGTTGAGATGGAAATGGACACATCAGTGTCCCTGTCATCTGTTCATCTGAACAATGTCATCTGT    4779

TCCCCTATGCGGTAGCGACCTTTACTGTGTCAGAAATATATTTAAGAAAATGTTCTAAAAAAAAAAAAAGGGGCCGGC    4858
```

FIG. 5G

PF00357

PF00357

PF00092

Cys
Ngly
out
TM
ins.

1.7 ■ ■3.8

6.4

1 41 81 121 161 201 241 281 321 361 401 441 481 521 561 601 641 681 721 761 801 841 881 921 961 1001 1041 1081 1121 1161
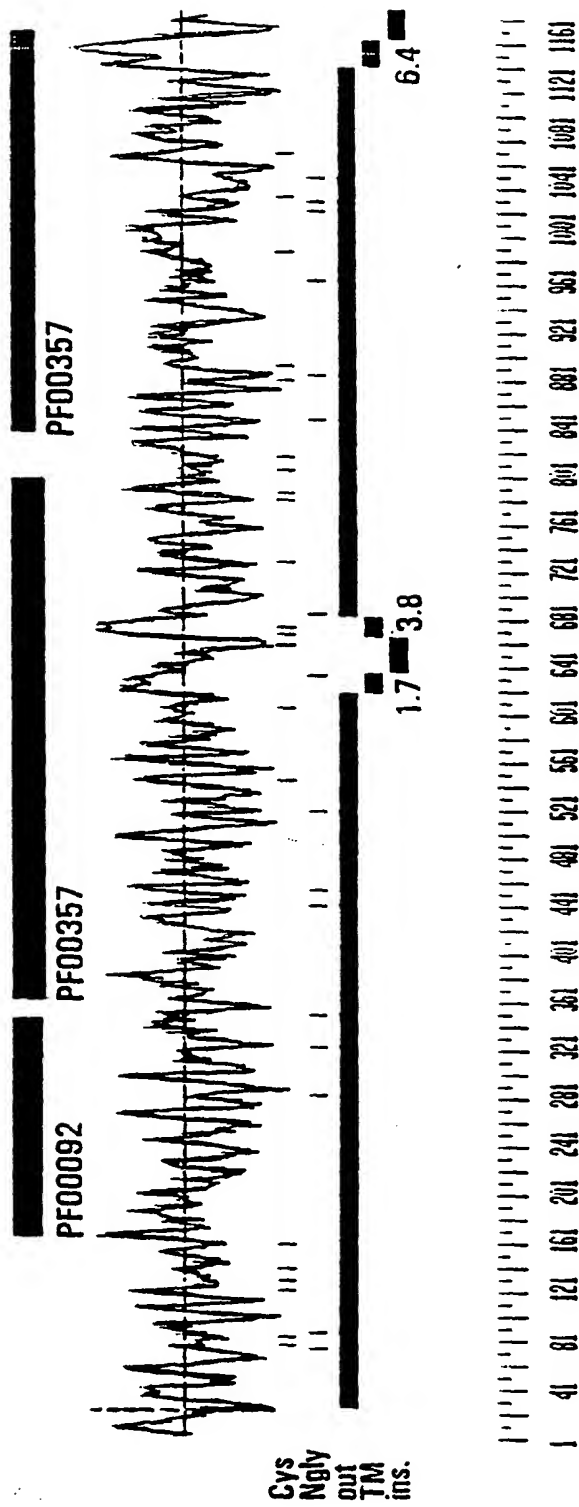
FIG. 6

```
              10        20        30        40        50        60        7
inputs GTCGACCCACGCGTCCGGGCGCCCCGCGCAGTGCAGGCTGCAGGCGCCGCGCCGAGGAGGCTGCCGCTC
       :
       G-------------------------------------------------------------------


              80        90       100       110       120       130       14
inputs GGCTTGCCGCCCCCGCCGCCGCTGCACACCGGACCCAGCCGCCGTGCCGCGGGCCATGGACCTGCCC
       ::  :  ::::::::::::                          :: :::::::::::::  : :::
       -------CGACACCCGCCGCCGC--------------------------GCAGCGGGCCATGGACTTTCCC
                     10                                       20        30        4(


             150       160       170       180       190       200       2:
inputs GGGCCTGGTGGTGGCCTGGGCGCTCAGCCTGTGGCCAGGGTTCACGGACACCTTCAACATGGACACCA(
       ::::::  ::::::::::::.: :::::::: ::::::::.::::::::::::::::::::  :::: ::::
       GGGCCTCCTGGTGGCCTGGACCCTCAGCCTTTGGCCAGGATTCACGGACACCTTCAATATGGATACCA(
          50        60        70        80        90       100       11(


             220       230       240       250       260       270       2:
inputs AAGCCCCGGGTCATCCCTGGCTCCAGGACCGCCTTCTTTGGCTACACAGTGCAGCAGCACGACATCAG'
       :: :::::::::: ::::::  :: .. :::::::::::::::::::::::::::::::::  :: :::::
       AATCCCCGGGTCATTGCTGGTCCCAGTGCTGCCTTCTTTGGCTACACAGTACAGCAGCATGATATCAG'
          120       130       140       150       160       170       18i


             290       300       310       320       330       340       3:
inputs GCAATAAGTGGCTGGTCGTGGGCGCCCCCACTGGAAAACCAATGGCTACCAGAAGACGGGAGACGTGTAC
       :::: :::::: :::: ::::: ::::: :::::::::::: : :::::::::.:::::::::.::
       GCAAGAAGTGGTTGGTTGTGGGTGCCCCAATGGAAAACCAATGGCCATCAGAAGACAGGAGACGTATAC
          190       200       210       220       230       240       25·


             360       370       380       390       400       410       4·
inputs GTGTCCAGTGATCCACGGGAACTGCACCAAACTCAACCTGGGAAGGGTCACCCTGTCCAACGTGTCCG
       .:::::.::.: ::: :: :::::::::::.::::::::::: ::::::: ::::::::: ::::: :
       ATGTCCGGTAACCCAGGGCAACTGCACCAAGCTCAACCTGGGCAGGGTCACTCTGTCCAATGTGTCTG
          260       270       280       290       300       310       32


             430       440       450       460       470       480       4
inputs CGGAAAGACAACATGCGCCTCGGCCTTAGTCTCGCCACCAACCCCAAGGACAACAGCTTCCTGGCCTG
       ::::.::::::::::::::::::::::: :: :: :::::::::::::::::::::::::::::::::::
       AGGAAGGACAACATGCGCCTCGGCCTGAGCCTTGCCACCAACCCCAAGGACAACAGCTTCCTGGCCTG
          330       340       350       360       370       380       39


             500       510       520       530       540       550       5
inputs GCCCCCTCTGGTCTCATGAGTGTGGGAGCTCCTACTACACCACAGGGATGTGTTCAAGAGTCAACTCC
       ::::  ::::::::: :: :::::::::.:::::::::::::::::::.::  :::::  :::.:::::::::
       GCCCTCTCTGGTCGCACGAGTGTGGAAGCTCCTACTACACCACTGGCATGTGCTCACGGGTCAACTCC
          400       410       420       430       440       450       46


             570       580       590       600       610       620       6
inputs CTTCAGGTTCTCCAAGACCGTGGCCCCAGCTCTCCAAAGGTGCCAGACCTACATGGACATCGTCATTG
       ::::::.:::::  :::::  ::.:: ::.::::. ::.::::::::  :::::::::::::::::::::::
       CTTCAGATTCTCTAAGACGGTAGCTCCGGCACTTCAGAGGTGCCAGACTTACATGGACATCGTCATTG
          470       480       490       500       510       520       53


             640       650       660       670       680       690       7
inputs CTGGATGGCTCCAACAGCATCTACCCCTGGGTGGAGGTTCAGCACTTCCTCATCAACATCCTGAAAAA
       :.::::::::::::::::::::::::::::::::::::: :: .:::::::::::: ::::: :::::::
       TTAGATGGCTCCAACAGCATCTACCCCTGGGTGGAGGTCCAACACTTCCTCATCAATATCCTCAAAAA
          540       550       560       570       580       590       60
```

FIG. 7A

```
            710       720       730       740       750       760       770
inputs TTTACATTGGCCCAGGGCAGATCCAGGTTGGAGTTGTGCAGTATGGCGAAGATGTGGTGCATGAGTTTCA
       : :::::::::::: :: :::::::::::: :::.:.:: :::::::: :::::::: :: ::::::::: ::
       TCTACATTGGCCCCGGCCAGATCCAGGTCGGAATAGTCCAGTATGGAGAAGATGCCGTCCATGAGTTCCA
            610       620       630       640       650       660       670


            780       790       800       810       820       830       840
inputs CCTCAACGACTACAGGTCTGTAAAAGATGTGGTGGAAGCTGCCAGCCACATTGAGCAGAGAGGAGGAACA
       ::: :: ::::::::::::::::::::::::::::::::::: :::::::::::::::::::::::::.:::
       CCTTAATGACTACAGGTCTGTAAAAGATGTGGTGGAAGCCGCCAGCCACATTGAGCAGAGAGGAGGGACA
            680       690       700       710       720       730       740


            850       860       870       880       890       900       910
inputs GAGACCCGGACGGCATTTGGCATTGAATTTGCACGCTCAGAGGCTTTCCAGAAGGGTGGAAGGAAAGGAG
       ::::::::: ::::::::::::::::::::::::::::::::.::::::::::::::::::::::::.:::::.:
       GAGACCCGCACGGCATTTGGCATTGAATTTGCACGCTCGGAGGCTTTCCAGAAGGGTGGAAGAAAAGGGG
            750       760       770       780       790       800       810


            920       930       940       950       960       970       980
inputs CCAAGAAGGTGATGATTGTCATCACAGATGGGGAGTCCCACGACAGCCCAGACCTGGAGAAGGTGATCCA
       :::::::.::::::::::::::::::.:: ::::::::::::::::::::::::::::::::::::::::::::.
       CCAAGAAAGTGATGATTGTCATCACGGACGGGGAATCCCACGACAGCCCAGACCTGGAGAAGGTGATCCG
            820       830       840       850       860       870       880


            990      1000      1010      1020      1030      1040      1050
inputs GCAAAGCGAAAGAGACAACGTAACAAGATATGCGGTGGCCGTCCTGGGCTACTACAACCGCAGGGGGATC
       ::::.:::::.:..:::::::.:: ::::: :: :::::::: :::::::::::::::::::::::::::::::
       GCAGAGCGAGAAGGACAACGTGACCAGATACGCTGTGGCCGTTTTGGGCTACTACAACCGCAGGGGGATC
            890       900       910       920       930       940       950


           1060      1070      1080      1090      1100      1110      112C
inputs AATCCAGAAACTTTTCTAAATGAAATCAAATACATCGCCAGTGACCCTGATGACAAGCACTTCTTCAATG
       :::::::::.:::::::::::::::::::::::::::::::::::: :::::::: ::::::::::::::::: :
       AATCCAGAGACTTTTCTAAATGAAATCAAATACATCGCCAGCGACCCTGACGACAAGCACTTCTTCAACC
            960       970       980       990      1000      1010      1020


           1130      1140      1150      1160      1170      1180      1190
inputs TCACTGATGAGGCTGCCTTGAAGGACATTGTCGATGCCCTGGGGGACAGAATCTTCAGCCTGGAAGGCAC
       ::::.:::::::: ::: ::::::::::::::::::::::::::::::: :::::::: :::::::::::::::
       TCACAGATGAGGCGGCCCCTGAAGGACATTGTTGATGCCCTTGGGGACAGGATCTTCAGCTTGGAAGGCAC
            1030      1040      1050      1060      1070      1080      1090


           1200      1210      1220      1230      1240      1250      126C
inputs CAACAAGAACGAGACCTCCTTTGGGCTGGAGATGTCACAGACGGGCTTTTCCTCGCACGTGGTGGAGGAT
       :::::::::: ::::::: :::::::::::::::::::::.::::: :::::::: :::.::::::.::::::
       CAACAAGAATGAGACCTCTTTTGGGCTGGAGATGTCGCAGACTGGCTTTTCTTCACATGTGGTAGAGGAT
            1100      1110      1120      1130      1140      1150      1160


           1270      1280      1290      1300      1310      1320      133C
inputs GGGGTTCTGCTGGGAGCCGTCGGTGCCTATGACTGGAATGGAGCTGTGCTAAAGGAGACGAGTGCCGGGA
       :::.: :::::::::::: :: ::.::::::::::::::: :::.:: ::::::.::::::::::.::::: :: :
       GGGATCCTGCTGGGAGCTGTGGGAGCCTATGACTGGAACGGGGCGGTGCTGAAGGAGACAAGTGCTGGCA
            1170      1180      1190      1200      1210      1220      1230


           1340      1350      1360      1370      1380      1390      140C
inputs AGGTCATTCCTCTCCGCGAGTCCTACCTGAAAGAGTTCCCCGAGGAGCTCAAGAACCATGGTGCATACCT
       :::: ::::::::::.::: ::::::::::::: ::::::::.:: :::::::::::::::::::.:: :::::
       AGGTGATTCCTCACCGAGAGTCCTACCTTAAGGAGTTTCCAGAGGAGCTGAAGAACCATGCAGCATACCT
            1240      1250      1260      1270      1280      1290      1300


           1410      1420      1430      1440      1450      1460      147C
inputs GGGGTACACAGTCACATCGGTCGTGTCCTCCAGGCAGGGGCGAGTGTACGTGGCCGGAGCCCCCCGGTTC
       .:::::::::.:: ::.::::: :::::::::::::::::::::::.::::: ::::: :::::::::: :.:::
       AGGGTACACGGTGACGTCGGTTGTGTCCTCCAGGCAGGGGCGGGTGTATGTGGCTGGAGCCCCCCAGATTC
            1310      1320      1330      1340      1350      1360      1370
```

FIG. 7B

```
            1480       1490       1500       1510       1520       1530       1540
inputs AACCACACGGGCAAGGTCATCCTGTTCACCATGCACAACAACCGGAGCCTCACCATCCACCAGGCTATGC
       :::::  :: :::::::::::: ::..::::  :::::::::::::::::::::::::::::::::::::::::  : :
       AACCATACTGGCAAGGTCATTCTATTCAGCATGCACAACAACCGGAGCCTCACCATCCACCAGGCTCTTC
            1380       1390       1400       1410       1420       1430       1440

            1550       1560       1570       1580       1590       1600       1610
inputs GGGGCCAGCAGATAGGCTCTTACTTTGGGAGTGAAATCACCTCGGTGGACATCGACGGCGACGGCGTGAC
       ::::: :::::::::::: :::::::::::::: ::::::.:::::::.: :. :::.: :::::
       GGGGCGAGCAGATAGGCTCCTACTTTGGGAGTGAGATTACCTCAGTGGACGTCAATGATGACAGAGTGAC
            1450       1460       1470       1480       1490       1500       1510

            1620       1630       1640       1650       1660       1670       1680
inputs TGATGTCCTGCTGGTGGGCGCACCCATGTACTTCAACGAGGGCCGTGAGCGAGGCAAGGTGTACGTCTAT
       .::::: ::::::::::: :::::::::::::.:::::::::.::::::::::::::::::::::::: :::::
       AGATGTGCTGCTGGTGGGTGCACCCATGTACTTCAGCGAGGGCCGAGAGCGAGGCAAGGTGTATGTCTAC
            1520       1530       1540       1550       1560       1570       1580

            1690       1700       1710       1720       1730       1740       1750
inputs GAGCTGAGACAGAACCGGTTTGTTTATAACGGAACGCTAAAGGATTCACACAGTTACCAGAATGCCCGAT
       .: :::::::::::::::::::::::::::: :: :: :::::::: :::::::: :::::::::: :::::::
       AACCTGAGACAGAACCGGTTTGTTTATAATGGCACTCTGAAGGATTCCCACAGCTACCAGAACGCCCGGT
            1590       1600       1610       1620       1630       1640       1650

            1760       1770       1780       1790       1800       1810       1820
inputs TTGGGTCCTCCATTGCCTCAGTTCGAGACCTCAACCAGGATTCCTACAATGACGTGGTGGTGGGAGCCCC
       : ::::::: ::: ::::.:::.:::::::::::::.:::::::::::::::: ::::::::::::.:::::
       TCGGGTCCTGCATCGCCTCTGTTCAAGACCTCAACCAAGATTCCTACAATGATGTGGTGGTGGGGGCCCC
            1660       1670       1680       1690       1700       1710       1720

            1830       1840       1850       1860       1870       1880       189C
inputs CCTGGAGGACAACCACGCAGGAGCCATCTACATCTTCCACGGCTTCCGAGGCAGCATCCTGAAGACACCI
       ::::::::::::.::::. :::.:::::::::::::::: :::::::.:. ::. :::::::::: .::
       TCTGGAGGACAGCCACAGAGGGGGCCATCTACATCTTCCATGGCTTCCAAACCAATATCCTGAAGAAGCCC
            1730       1740       1750       1760       1770       1780       1790

            1900       1910       1920       1930       1940       1950       196C
inputs AAGCAGAGAATCACAGCCTCAGAGCTGGCTACCGGCCTCCAGTATTTTGGCTGCAGCATCCACGGGCAAT
       :.:::::::: :::::::::.::::::: :  ::::: ::: :::::::::::::::::::::::::::::::::
       ATGCAGAGAATAACAGCCTCGGAGCTGGCTCCTGGCCTGCAGCATTTTGGCTGCAGCATCCACGGACAAC
            1800       1810       1820       1830       1840       1850       1860

            1970       1980       1990       2000       2010       2020       203C
inputs TGGACCTCAATGAGGATGGGCTCATCGACCTGGCAGTGGGAGCCCCTTGGCAACGCTGTGATTCTGTGGTC
       :::::::::::::::::::::::::: .: :::::.:::::::: ::::: ::::::::::::.:: :::::: :
       TGGACCTCAATGAGGATGGGCTTGTGGACCTAGCAGTGGGCGCCCTGGGCAACGCTGTGGTTTTGTGGGC
            1870       1880       1890       1900       1910       1920       1930

            2040       2050       2060       2070       2080       2090       210C
inputs CCGCCCAGTGGTTCAGATCAATGCCAGCCTCCACTTTGAGCCATCCAAGATCAACATCTTCCACAGAGAC
       :::::::.::.:::::.:::::  ::::::::::::::::::::::.::::::::::::::::::::..:I:
       CCGCCCTGTAGTTCAAATCAACGCCAGCCTCCACTTTGAGCCTTCCAAGATCAACATCTTCCACAAGGAC
            1940       1950       1960       1970       1980       1990       2000

            2110       2120       2130       2140       2150       2160       217C
inputs TGCAAGCGCAGTGGCAGGGATGCCACCTGCCTGGCCGCCTTCCTCTGCTTCACGCCCATCTTCCTGGCAC
       :::::::::::.::::::::::::::::::::::::::::::::::::::::::: .::::::::::::::::::
       TGCAAGCGCAATGGCAGGGATGCCACCTGCCTGGCCGCCTTCCTCTGCTTCATACCCATCTTCCTGGCAC
            2010       2020       2030       2040       2050       2060       2070

            2180       2190       2200       2210       2220       2230       224C
inputs CCCATTTCCAAACAACAACTGTTGGCATCAGATACAACGCCACCATGGATGAGAGGCGGTATACACCGAC
       :::: :::::::::::.::::: :: :::::::::.::::: :: ::::::::::::::::::::::::: .::: :
       CCCACTTCCAAACAGCAACCGTCGGCATCAGGTACAATGCAACCATGGATGAGAGGCGGTATATGCCGCC
            2080       2090       2100       2110       2120       2130       2140
```

FIG. 7C

```
            2250      2260      2270      2280      2290      2300      2310
inputs GGCCCACCTGGACGAGGGCGGGGACCGATTCACCAACAGAGCCGTACTGCTCTCCTCCGGCCAGGAGCTC
       ::: ::  :::::::::::::.::::..:::::::::::::::::: :::::::::::::::::::::.:.:
       GGCACATCTGGACGAGGGCGGAGACCAGTTCACCAACAGAGCCGTCCTGCTCTCCTCTGGCCAGGAACAC
            2150      2160      2170      2180      2190      2200      2210

            2320      2330      2340      2350      2360      2370      2380
inputs TGTGAGCGGATCAACTTCCATGTCCTGGACACTGCTGACTACGTGAAGCCAGTGACCTTCTCAGTCGAGT
       ::: :. :::::::::::::::::::::::::::::::.:::::::::::::::::::.::::::: :: ::::
       TGTCAAAGGATCAACTTCCATGTCCTGGACACTGCAGACTACGTGAAGCCAGTGGCCTTCTCCGTGGAGT
            2220      2230      2240      2250      2260      2270      2280

            2390      2400      2410      2420      2430      2440      2450
inputs ATTCCCTGGAGGACCCTGACCATGGCCCCATGCTGGACGACGGCTGGCCCACCACTCTCAGAGTCTCGGT
       : :::::.:::::::::::::::·: : :::::::::::::.:::::::::::::: :: :::::::: ::::
       ACTCCCTAGAGGACCCTGACAACGGCCCCATGCTGGACAACGGCTGGCCCACTACGCTCAGAGTGTCGGT
            2290      2300      2310      2320      2330      2340      2350

            2460      2470      2480      2490      2500      2510      2520
inputs GCCCTTCTGGAACGGCTGCAATGAGGATGAGCACTGTGTCCCTGACCTTGTGTTGGATGCCCGGAGTGAC
       :::::::::: ::::: ::::::::::::.::::::::::::::::::::::::. :::::: ::::::::::
       GCCCTTCTGGAATGGCTGTAATGAGGATGAACACTGTGTCCCTGACCTTGTACTGGATGCTCGGAGTGAC
            2360      2370      2380      2390      2400      2410      2420

            2530      2540      2550      2560      2570      2580      2590
inputs CTGCCCACGGCCATGGAGTACTGCCAGAGGGTGCTGAGGAAGCCTGCGCAGGACTGCTCCGCATACACGC
       :::::::: :::::::::::::::::::.::::::::: ::::::::.::::::::::::::::.  ::::: :
       CTGCCCACTGCCATGGAGTACTGCCAGCGGGTGCTGGGGAGGCCGGCGCAGGACTGCTCCAGCTACACCC
            2430      2440      2450      2460      2470      2480      2490

            2600      2610      2620      2630      2640      2650      2660
inputs TGTCCTTCGACACCACAGTCTTCATCATAGAGAGCACACGCCAGCGAGTGGCGGTGGAGGCCACACTGGA
       ::::::::::::::::.:::::::::::::::::::::.::::.  ::.::::::.::::::::::::.:::::
       TGTCCTTCGACACCACTGTCTTCATCATAGAGAGCACGCGCCGCCGGGTGGCAGTGGAGGCCACGCTGGA
            2500      2510      2520      2530      2540      2550      2560

            2670      2680      2690      2700      2710      2720      2730
inputs GAACAGGGGCGAGAACGCCTACAGCACGGTCCTAAATATCTCGCAGTCAGCAAACCTGCAGTTTGCCAGC
       :::::::.::  :::::  :: :::::::.: :::::  ::  ::::: :::::::  .::::: :::::::::::
       GAACAGAGGAGAGAATGCTTACAGCGCCGTCCTCAACATCTCCCAGTCAGAGAACCTTCAGTTTGCCAGC
            2570      2580      2590      2600      2610      2620      2630

            2740      2750      2760      2770      2780      2790      2800
inputs TTGATCCAGAAGGAGGACTCAGACGGTAGCATTGAGTGTGTGAACGAGGAGAGGAGGCTCCAGAAGCAAG
       :::::::::::::::: ::::::::.. ::::: ::::::::::::::::::::::: :::: :: ::: :::
       CTGATCCAGAAGGATGACTCAGACAACAGCATCGAGTGTGTGAACGAGGAGAGGCGGCTTCACAAGAAAG
            2640      2650      2660      2670      2680      2690      2700

            2810      2820      2830      2840      2850      2860      2870
inputs TCTGCAACGTCAGCTATCCCTTCTTCCGGGCCAAGGCCAAGGTGGCTTTCCGTCTTGATTTTGAGTTCAG
       :::::::::::::::::: :::::::: :.:::::::::::::::::::::::::::: ::::: ::::::::
       TCTGCAACGTCAGCTACCCCTTCTTCAGAGCCAAGGCCAAGGTGGCTTTCCGTCTGGATTTCGAGTTCAG
            2710·     2720      2730      2740      2750      2760      2770

            2880      2890      2900      2910      2920      2930      2940
inputs CAAATCCATCTTCCTACACCACCTGGAGATCGAGCTCGCTGCAGGCAGTGACAGTAATGAGCGGGACAGC
       :::.:: .: :: ::.:::::: ::  ::::: : :: ::: :::::::::::: : :::::..::::::::
       CAAGTCTGTGTTTCTGCACCATCTTCAGATCCATCTGGGTGCCGGCAGTGACAGTCACGAGCAAGACAGC
            2780      2790      2800      2810      2820      2830      2840
```

FIG. 7D

```
           2950      2960      2970      2980      2990      3000      3010
inputs ACCAAGGAAGACAACGTGGCCCCCTTACGCTTCCACCTCAAATACGAGGCTGACGTCCTCTTCACCAGGA
       :::. .:: :: :::. ::::: :  :.::::::::::::::::: ::.::.:::::::::: :::::..:
       ACCGCAGACGATAACACGGCCCTCCTGCGCTTCCACCTCAAATATGAAGCAGACGTCCTCTTTACCAGAA
           2850      2860      2870      2880      2890      2900      2910


           3020      3030      3040      3050      3060      3070      3080
inputs GCAGCAGCCTGAGCCACTACGAGGTCAAGCCCAACAGCTCGCTGGAGAGATACGATGGTATCGGGCCTCC
       :::::::::::::::::::::.::::::::::: : :::::::::.:: ::::: :: ::::: :::::.:::::
       GCAGCAGCCTGAGCCACTTCGAGGTCAAGGCAAACAGCTCACTTGAGAGCTATGATGGCATCGGACCTCC
           2920      2930      2940      2950      2960      2970      2980


           3090      3100      3110      3120      3130      3140      315C
inputs CTTCAGCTGCATCTTCAGGATCCAGAACTTGGGCTTGTTCCCCATCCACGGGATGATGATGAAGATCACC
       :::::.::: .: :::::.: ::::: ::::::: ::::::::::::.::::::::.:::::::::::::::
       GTTCAACTGTGTTTTCAAGGTGCAGAACCTGGGCTTTTTTCCCCATCCACGGGGTGATGATGAAGATCACT
           2990      3000      3010      3020      3030      3040      3050


           3160      3170      3180      3190      3200      3210      3220
inputs ATTCCCATCGCCACCAGGAGCGGCAACCGCCTACTGAAGCTGAGGGACTTCCTCACGGACGAGGCGAACA
       .: :::::::::::::::::.: :::::::::::::::::::.:::::.: ::::::::::::: :::: ::: ::: ::::
       GTGCCCATCGCCACCAGGGGTGGCAACCGCCTGCTGATGTTGAGGGACTTCTTCACCGACCAGGGCAACA
           3060      3070      3080      3090      3100      3110      3120


           3230      3240      3250      3260      3270      3280      3290
inputs CGTCCTGTAACATCTGGGGCAATAGCACTGAGTACCGGCCCACCCCAGTGGAGGAAGACTTGCGTCGTGC
       :.:::::::::::::::: :: :::::.::::::  ::  :::::. ::::::::: :: : :.:::
       CATCCTGTAACATCTGGGGGAACAGCACAGAGTACAGGAGTACCCCAACTGAGGAAGATCTGAGCCATGC
           3130      3140      3150      3160      3170      3180      3190


           3300      3310      3320      3330      3340      3350      336C
inputs TCCACAGCTGAATCACAGCAACTCTGATGTCGTCTCCATCAACTGCAATATACGGCTGGTCCCCAACCAC
       :::::: ::::::::::::: :: ::::::::::.::::: :. :::::: ::::::::.:::
       CCCACAGAGGAATCACAGCAACTCCGACGTGGTCTCCATCATCTGCAACCTGAGGCTGGCCCCCAGCCAC
           3200      3210      3220      3230      3240      3250      3260


           3370      3380      3390      3400      3410      3420      343C
inputs GAAATCAATTTCCATCTACTGGGGAACCTGTGGTTGAGGTCCCTAAAAGCACTCAAGTACAAATCCATGA
       ::.: ::. ::: : ::. ::::.:::::::::: ::: .:: ::.:::::::::::::::::::..:: :::
       GAGACCAGCTTCTACCTGGTGGGAAACCTGTGGCTGACATCTCTGAAAGCACTCAAGTACAGGTCTCTGA
           3270      3280      3290      3300      3310      3320      3330


           3440      3450      3460      3470      3480      3490      3500
inputs AAATCATGGTCAACGCAGCCTTGCAGAGGCAGTTCCACAGCCCCTTCATCTTCCGTGAGGAGGATCCCAC
       :..:::: .::::::::: :::::::::::::::::::::::::::::::::::::::::::.:::::.:: :: ::
       AGATCACAGTCAACGCCGCCTTGCAGAGGCAGTTCCACAGCCCCTTCATCTTCCGAGAGGAAGACCCTAG
           3340      3350      3360      3370      3380      3390      3400


           3510      3520      3530      3540      3550      3560      3570
inputs CCGCCAGATCGTGTTTGAGATCTCCAAGCAAGAGGACTGGCAGGTCCCCATCTGGATCATTGTAGGCAGC
       :::::::::.. .::::::::::::::::::::::::::.:::::::::::::::::::::: ::.:::::::
       CCGCCAGGTCACATTTGAGATCTCCAAGCAAGAAGACTGGCAGGTCCCCATCTGGATCATCGTGGGCAGC
           3410      3420      3430      3440      3450      3460      3470


           3580      3590      3600      3610      3620      3630      3640
inputs ACCCTGGGGGGCCTCCTACTGCTGGCCCTGCTGGTCCTGGCACTGTGGAAGCTCGGCTTCTTTAGAAGTG
       :: :::::::::::: :.:::::::::::::::::::::::::::::::::::::::::::: :: ::::.:::::
       ACTCTGGGGGGCCTCTTGCTGCTGGCCCTGCTGGTCCTGGCACTGTGGAAGCTCGGTTTTTTTAAAAGTG
           3480      3490      3500      3510      3520      3530      3540


           3650      3660      3670      3680      3690      3700      3710
inputs CCAGGCGCAGGAGGGAGCCTGGTCTGGACCCCACCCCCAAAGTGCTGGAGTGAGGCTCCAGAGGAGACTT
       :::::.:::::.:.:::::::::::. :::.:::. :::::::.::: .:::::: :::::::::::..:::
       CCAAGCGCAAGAGGGAACCTGGCCTAGGCCCCATCCCCAAAGAGCTCAAGTGAGGATCCAGAGGAAGCTT
           3550      3560      3570      3580      3590      3600      3610
```

FIG. 7E

```
              3720       3730       3740       3750       3760       3770      378
inputs TGAGTTGATGGGGGCCAGGACACCAGTCCAGGTAGTGTTGAGACCCAGGCCTGTGGCCCCACCGAGCTC
       .:::::.:::::::: : :::::::::::.:: .:::: ::.:::.::: :.::::: :: :::::::
       CAAGTTGTTGGGGGCCCG-ACACCAGTCCGGG-GGTGTGCAGGCCCGGGC-TATGGCCTCA--GAGCTC
              3620       3630       3640       3650       3660       3670

              3790       3800       3810       3820       3830       3840      385
inputs AGCGGAGAGGAAGCCAGCTGGCTTTGCACTTGACCTCATCTCCCGAGCAATGGCGCCTGCTCCCTCCAC
       :::..::::::: .::: . :.:::::.:: :::: ::::::::: :::::::::::::. ::: :: :: ..:
       AGCAAAGAGGACACCCATGGACTTTACA-TTGAGTTCATCTCTGGAGCAATGGCAACTGATCGCTGTGC
       3680       3690       3700       3710       3720       3730       3740

              3860       3870       3880       3890       3900       3910      392
inputs ATGGAACTCAAGCTGGTTTTAAGTGGAACTGCCCTACTGGGAGACTGGGACACCTTTAACACAGACCCC
       :::::: ::: ::: :: :: ..: .::: : :::::.:..: :..:: :..::. :: :::: :::
       ATGGACCTCCAGCCCGTGTTCGGAGTGACTTCTCTACTAGAGGCCAAGACCTCCAGTA---CAGATCCC
       3750       3760       3770       3780       3790       3800       3810

              3930       3940       3950       3960       3970       3980
inputs AGGGATTTAAAGGGACACCCCTACACACACCCAGGCCCACGCCAAGGCCT-CCCTCA-GGCTCTGTGGA
       :..:: :.:::::::: :::: :: ::.:::::: ..::::: :: ::::::::::::
       AAGGACTAAAAGGGA---CCCT-------CCAAGACCCAC-CGGTGGCCTTCCCCCAAGGCTCTGTGGA
       3820       3830          3840       3850       3860       387C

       3990       4000       4010       4020       4030       4040       4050
inputs GGCATTTGCTGCCCCAGCTACTAAGGTGCTAGGAATTCGTAATCATCCCCATCCTCCAGAGAAACCCAC
       :: :: :::::::::. ::::::::::::::::..:: ::..::.: :::: :::: ..:::::: ::.
       GGAATCTGCTGCCCCAAACACTAAGGTGCTAGGGTTTTGTTGTCTTGCCCA-CCTCTGAAGAAACTCA
       3880       3890       3900       3910       3920       3930       394

       4060       4070       4080       4090       4100       4110       4120
inputs GAGGAAGACTGTAAATACGAACCCAATCTGCACACTCCAGGCCTCTAGTTCCAGAAGGATCCAAGACA
       : ::::: :: :::::::: :: :..:: ::..:: ::.. :.:::.
       G-GGAAGCTAGCAAATACGAGCCCACTGTGCACGC-CCTA---TCTGGA-CCTATAGCTTC-AGGACTC
       3950       3960       3970         3980       3990       4000

       4130       4140       4150       4160       4170       4180       4190
inputs ACAGATCTGAATTCTGCCCTTTTCTCTCACCCATCCCACCCCTCCA-TTGGCTCCCAAGTCACACCCAC
       : .: : :.: :. ::: : : ::::: : :::: ::. :::: ::::::. :::: :::
       ATGGCCCGGGAGCATAACCTCCGC-CGCACCC--CTTACCCTTGCTCTTGGTTCCCAAACCACAGCCA-
       4010       4020       4030       4040       4050       4060

       4200       4210       4220       4230       4240       4250       4260
inputs CCCTTCCCCATAGATAGGCCCCTGGGGCTCCTGAAGAATGAACCCAAGAGCAAGGGCTTGATGGTGACA
       :::::: ::: :::::::.::::...:. :: :.::.::: :: .::. .. :::::::. :: ::
       CCCTTCTCCACAGATAGACCCCAAAGA-TCTCTATGAGTGACCCAGAGTTGTGTGGCTTGATA-TGCCA
       4070       4080       4090       4100       4110       4120       4130

       4270       4280       4290       4300       4310       4320       4330
inputs CTGCAAGCCAGGGATGAAGAAAGACTCTGAGATGTGGAGACTGATGGCCAGGCAAGTGGGACCAGGATA
       :::::..::::: :::.::.: : :.:.: . :::: : ::.: ::..: ::.: ::..:
       CTGCAGGCCAGTGATGGAAACACATCCAGGGAC--TGTGCCAAATGAACTGG---AGGGAATCAGGATC
       4140       4150       4160       4170       4180       4190       4200

       4340       4350       4360       4370       4380       4390       4400
inputs TGGACGCTGTCCTGAGATGAGAG-GTAGCCGGGCTCTGCACCCACGTGCATTCACATTGACCGCAACTC
       :::. .: :::..: :: :::. : ::: ::::::: ::: .: :.:::.: :: :
       TGGGTACCTTCCAAACATCAGAAAGCAGCAGGGCTCT----CCA---------GCAGTAACCACCACCC
           4210       4220       4230       4240              4250
```

FIG. 7F

```
          4410       4420       4430       4440       4450       4460       4470
inputs CACATTCC--CCCACCAGCTGCAGCCCCTTGCTCTCAGCTGCCAACCCTCCCGGGTCACTTTTGTTC(
       :.::.::: : :.:.::::: :. :::: : ::::: :::: :.: :.::::: :::
          CTCAATCCTGCTCTCCAGC--CT-CCCC-----CGCAGCT-----CCTCTCAG-CCGCTTTGGTTG(
        4260       4270       4280       4290            4300       4310


          4480       4490       4500       4510       4520       4530
inputs GGTACCTCATG---GGAAGC---ATGTGGATGACACAATCCCTGGGGCTGTGCATTCCCAC--GTCT
       :: : :::.:: ::.::: : :.::::: : ::. :: : :::: : :.:.::: .: ::::
          GGCAACTCGTGCTAGGGAGCCTGAGGAGGATGATAAAAATCC-GTGGCTCTCCGTACCCTCTTGTCT(
        4320       4330       4340       4350       4360       4370       4380


          4540       4550       4560       4570       4580       4590       4600
inputs TGCTGCAGCCTGCCCCTAGACATGGACGCACTGGCCTGGCTGCAGCTGGGCAGCAGGGGTAGGGGTA(
       ..::::: :: :::::.: .: ::::: :: :: :.: :...
          AACTGCAGCTTGTCCCTGG---TAGATGCACTG---TGATAG----GGAGCCTCTGTTTCTTACATC(
        4390       4400       4410            4420       4430       4440


          4610       4620       4630       4640       4650       4660       4670
inputs AGCCTCCCCTCCCTGTATCACCCCCTCCCTACACACACACACACACACACACACACACACACT-GCCTCC
       : ::: : :.: :. . .::: : :.: :::::::::::::::::::::::::::::.:: .:.:::
          A-CCTACACACACACACACACACAC-ACACACACACACACACACACACACACGCACGCACCTCC
        4450       4460       4470       4480       4490       4500      451


          4680       4690       4700       4710       4720       4730       4740
inputs TCCTTCCCTCATGCCCGC-CAGTGCACAGGGAAGGGCTTGGCCAGCGCTGTTGAGGGGTCCCCTCTGG
       .::.:: :.:: : :: ::.: : :.::: ::::::: :.::: :: :::.:
          ACCATCA-GCGTGTGGGAGAAGGGCTCCTGTAGGGG--------GCGCTGTGGGGGG---CCTTCTAG
        4520       4530       4540               4550       4560


          4750       4760       4770       4780       4790       4800       4810
inputs ATGCACTGAATAAAGCACGTGCAAGGACTCCCGGAGCCTGTGCAGCCTTGGTGGCAAATATCTCATCT
       :::::::::: ::::: :.:::::: :: : ::.::: . :.::: .:.::::: : :::::
          ATGCACTGAACAAAGCCCATGCAAGGCCCCTGGGAACCTTA--AACCTCAGAGGCAGATAG-GCATCT(
        4570       4580       4590       4600       4610       4620       4630


          4820       4830       4840       4850       4860       4870       4880
inputs CCGGCCCCCAGGACAAGTGGTATGACCAGTGATAATGCCCCAAGGACAAGGGGCGTGCCTGGCGCCCA(
       :::::: .... ..: ::::: :::::::::::::::::::.:::: :. :::::::
          CCGGCC-----AGCT-----------CTGGGATAA-GCCCCAAGGACAAGGGGCATGCCGGATGCCCA(
        4640                              4650       4660       4670       4680


          4890       4900       4910       4920       4930       4940
inputs GGAGTAATTTATGCCTTAGTCTTGTTTTGAGGTAGAAATG----CAAGGGGGACACATGAAAGGCATC/
       .:::::::::::::::::::. ::: : :::::.:.::::: ::. .: :.: : ::. ::::.
          AGAGTAATTTATGCCTTAACCTTTTGTTGAGATGGAAATGGACACATCAGTGTCCCCTGT----CATC1
        4690       4700       4710       4720       4730       4740       4750


          4950       4960       4970       4980       4990       5000     501
inputs TCCCC-CTGTGCATAGT--ACGA----CCTTTACTGT-CGTATTTTTGAAAAATTAAAAATACAGTGTT
       : : :::.:::: .: .. .: . ::.::: :: :: ::: ::. ... : .:::.:.. ...:.
          TTTCATCTGAACAATGTCATCTGTTCCCCTATGCGGTACGACCTTTACTGTCAGAAATATATTTAAGAA
        4760       4770       4780       4790       4800       4810       4820


             5020       5030       5040
inputs AAAAAC-AAAAAAAAAAAAAAAAAGGGCGGCCGC
       :.....: :::::::::::::::::::::::::::
          ATGTTCTAAAAAAAAAAAAAAAAAAGGGCGGCCGC
        4830       4840       4850
```

FIG. 7G

```
              10        20        30        40        50        60        7(
inputs MDLPRGLVVAWALSLWPGFTDTFNMDTRKPRVIPGSRTAFFGYTVQQHDISGNKWLVVGAPLETNGYQK
       ::.::::.::::::::::::::::::::.:::..:. .:::::::::::::::::.:::::::.:::: ::
       MDFPRGLLVAWTLSLWPGFTDTFNMDTRNPRVIAGPSAAFFGYTVQQHDISGKKWLVVGAPMETNGHQK
              10        20        30        40        50        60        7(

              80        90       100       110       120       130       14:
inputs GDVYKCPVIHGNCTKLNLGRVTLSNVSERKDNMRLGLSLATNPKDNSFLACSPLWSHECGSSYYTTGMC.
       ::::::::..:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       GDVYKCPVTQGNCTKLNLGRVTLSNVSERKDNMRLGLSLATNPKDNSFLACSPLWSHECGSSYYTTGMC:
              80        90       100       110       120       130       14(

             150       160       170       180       190       200       21(
inputs RVNSNFRFSKTVAPALQRCQTYMDIVIVLDGSNSIYPWVEVQHFLINILKKFYIGPGQIQVGVVQYGED
       :::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::.::::.:
       RVNSNFRFSKTVAPALQRCQTYMDIVIVLDGSNSIYPWVEVQHFLINILKKFYIGPGQIQVGIVQYGED:
             150       160       170       180       190       200       21:

             220       230       240       250       260       270       28:
inputs VHEFHLNDYRSVKDVVEAASHIEQRGGTETRTAFGIEFARSEAFQKGGRKGAKKVMIVITDGESHDSPD:
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       VHEFHLNDYRSVKDVVEAASHIEQRGGTETRTAFGIEFARSEAFQKGGRKGAKKVMIVITDGESHDSPD
             220       230       240       250       260       270       28

             290       300       310       320       330       340       35
inputs EKVIQQSERDNVTRYAVAVLGYYNRRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIVDALGDRI
       ::::.::::.:::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
       EKVIRQSEKDNVTRYAVAVLGYYNRRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIVDALGDRI:
             290       300       310       320       330       340       35

             360       370       380       390       400       410       42
inputs SLEGTNKNETSFGLEMSQTGFSSHVVEDGVLLGAVGAYDWNGAVLKETSAGKVIPLRESYLKEFPEELK
       :::::::::::::::::::::::::::::::::.::::::::::::::::::::::::: :::::::::::
       SLEGTNKNETSFGLEMSQTGFSSHVVEDGILLGAVGAYDWNGAVLKETSAGKVIPHRESYLKEFPEELK
             360       370       380       390       400       410       42

             430       440       450       460       470       480       49
inputs HGAYLGYTVTSVVSSRQGRVYVAGAPRFNHTGKVILFTMHNNRSLTIHQAMRGQQIGSYFGSEITSVDI
       :.:::::::::::::::::::::::::::::::::::.:::::::::::::::.::.:::::::::::::.
       HAAYLGYTVTSVVSSRQGRVYVAGAPRFNHTGKVILFSMHNNRSLTIHQALRGEQIGSYFGSEITSVDV
             430       440       450       460       470       480       49

             500       510       520       530       540       550       56
inputs GDGVTDVLLVGAPMYFNEGRERGKVYVYELRQNRFVYNGTLKDSHSYQNARFGSSIASVRDLNQDSYND
       .: :::::::::::::::.:::::::::::.:::::::::::::::::::::::::.:: ::.:::::::
       DDRVTDVLLVGAPMYFSEGRERGKVYVYNLRQNRFVYNGTLKDSHSYQNARFGSCIASVQDLNQDSYND
             500       510       520       530       540       550       56

             570       580       590       600       610       620       63
inputs VVGAPLEDNHAGAIYIFHGFRGSILKTPKQRITASELATGLQYFGCSIHGQLDLNEDGLIDLAVGALGN
       :::::::::.: :::::::::.:: .::: :.::::::::: ::: :::::::::::::::::.:::::::
       VVGAPLEDSHRGAIYIFHGFQTNILKKPMQRITASELAPGLQHFGCSIHGQLDLNEDGLVDLAVGALGN
             570       580       590       600       610       620       63

             640       650       660       670       680       690       70
inputs VILWSRPVVQINASLHFEPSKINIFHRDCKRSGRDATCLAAFLCFTPIFLAPHFQTTTVGIRYNATMDE
       :..::.::::::::::::::::::::::.::::.:::::::::::::.:::::::::::::.::::::::::
       VVLWARPVVQINASLHFEPSKINIFHKDCKRNGRDATCLAAFLCFIPIFLAPHFQTATVGIRYNATMDE
            ·640       650       660       670       680       690       70
```

FIG. 8A

```
                 710        720        730        740        750        760        77
inputs  RYTPRAHLDEGGDRFTNRAVLLSSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGPMLDDGWPT
        ::  :::::::::::.:::::::::::::  :..:::::::::::::::::.:::::::::::::.:::::.:::
        RYMPRAHLDEGGDQFTNRAVLLSSGQEHCQRINFHVLDTADYVKPVAFSVEYSLEDPDNGPMLDNGWPT
                 710        720        730        740        750        760        77

                 780        790        800        810        820        830        84
inputs  LRVSVPFWNGCNEDEHCVPDLVLDARSDLPTAMEYCQRVLRKPAQDCSAYTLSFDTTVFIIESTRQRVA
        :::::::::::::::::::::::::::::::::::::::::::: .::::::.::::::::::::::::::::
        LRVSVPFWNGCNEDEHCVPDLVLDARSDLPTAMEYCQRVLGRPAQDCSSYTLSFDTTVFIIESTRRRVA
                 780        790        800        810        820        830        84

                 850        860        870        880        890        900        91
inputs  EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRLQKQVCNVSYPFFRAKAKVAFR
        ::::::::.:::::::.:::::::.:.::::::::::.:.:.::::::::::::::.:..::::::::::::
        EATLENRGENAYSAVLNISQSENLQFASLIQKDDSDNSIECVNEERRLHKKVCNVSYPFFRAKAKVAFR
                 850        860        870        880        890        900        91

                 920        930        940        950        960        970        98
inputs  DFEFSKSIFLHHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLFTRSSSLSHYEVKPNSSLERY
        :::::::.:::::.:  :.:::::.:.::: .::.:  :::::::::::::::::::::::.:::.::::: :
        DFEFSKSVFLHHLQIHLGAGSDSHEQDSTADDNTALLRFHLKYEADVLFTRSSSLSHFEVKANSSLESY
                 920        930        940        950        960        970        98

                 990       1000       1010       1020       1030       1040       105
inputs  GIGPPFSCIFRIQNLGLFPIHGMMMKITIPIATRSGNRLLKLRDFLTDEANTSCNIWGNSTEYRPTPVE
        :::::::.:::.:::::::.::::.::::.:::::::.:::::.:::.:::.:::::::::::::::.::.:
        GIGPPFNCVFKVQNLGFFPIHGVMMKITVPIATRGGNRLLMLRDFFTDQGNTSCNIWGNSTEYRSTPTE
                 990       1000       1010       1020       1030       1040       105

                1060       1070       1080       1090       1100       1110       112
inputs  DLRRAPQLNHSNSDVVSINCNIRLVPNQEINFHLLGNLWLRSLKALKYKSMKIMVNAALQRQFHSPFIF
        ::  .:::  :::::::::::::  ::.::.:.::..:  :.:::::  ::::::::.::  ::::::::::::
        DLSHAPQRNHSNSDVVSIICNLRLAPSQETSFYLVGNLWLTSLKALKYRSLKITVNAALQRQFHSPFIF
                1060       1070       1080       1090       1100       1110       112

                1130       1140       1150       1160       1170       1180
inputs  EEDPSRQIVFEISKQEDWQVPIWIIVGSTLGGLLLLALLVLALWKLGFFRSARRRREPGLDPTPKVLE
        :::::::..:::::::::::::::::::::::::::::::::::::::::::::::::..:.:.:::::..:: :
        EEDPSRQVTFEISKQEDWQVPIWIIVGSTLGGLLLLALLVLALWKLGFFKSAKRKREPGLGPIPKELK
                1130       1140       1150       1160       1170       1180
```

FIG. 8B

```
int_alpha_rpt: domain 1 of 5, from 37 to 90: score 29.5, E = 7.6e-05
                    *-pgsyFGysvagvgDlngDksGypDlllvGAPr........GaVYvyf
                       ++FGy+v    +D+      G  + l+vGAP+++++ +++G VY+++
          T259   37    RTAFFGYTVQ-QHDIS----GNKW-LVVGAPLetngyqktGDVYKCP 77


                    gssnsgggrciplqnls<-*
                    +       g+c+ l    +
          T259   78 VI----HGNCTKLNLGR      90
```

Figure 9A

```
>int_alpha_rpt: domain 2 of 5, from 421 to 472: score 19.2, E = 0.097
                    *-pgsyFGysvagvgDlngDksGypDlllvGAPr....GaVYvyfgssn
                       +g+y+Gy+v++v+              ++++GAPr +++G+V ++  +
          T259   421    HGAYLGYTVTSVVSS-----RQGRVYVAGAPRfnhtGKVILFTMHN- 461


                    sgggrciplqnls<-*
                    +++++  q  +
          T259   462 --NRSLTIHQAMR      472
```

Figure 9B

```
>int_alpha_rpt: domain 3 of 5, from 476 to 532: score 48.4, E = 1.5e-10
                    *-pgsyFGysvagvgDlngDksGypDlllvGAPr........GaVYvyf
                       +gsyFG+ ++  +D++gD  G +D+llvGAP+  +++++++G+VYvy
          T259   476    IGSYFGSEIT-SVDIDGD--GVTDVLLVGAPMyfnegrerGKVYVYE 519


                    gssnsgggrciplqnls<-*
                    +       +r+ +  +l
          T259   520 LR----QNRFVYNGTLK      532
```

Figure 9C

```
>int_alpha_rpt: domain 4 of 5, from 538 to 593: score 62.6, E = 8.4e-15
               *-pgsyFGysvagvgDlngDksGypDlllvGAPr.....GaVYvyfgss
                 ++++FG+s+a+v+Dln+D  +y+D ++vGAP+++++ Ga+Y+++g +
      T259  538  QNARFGSSIASVRDLNQD--SYND-VVVGAPLednhaGAIYIFHGFR 581

               nsgggrciplqnls<-*
                 +   +   +p q ++
      T259  582 G--SILKTPKQRIT    593
```

Figure 9D

```
>int_alpha_rpt: domain 5 of 5, from 600 to 654: score 16.5, E = 0.57
               *-pgsyFGysvagvgDlngDksGypDlllvGAPr.....GaVYvyfgss
                 yFG s+ g  Dln D   G   D l+vGA ++        +v+++++
      T259  600  GLQYFGCSIHGQLDLNED--GLID-LAVGALGnavilWSRPVVQINA 643

               nsgggrciplqnls<-*
                 +++++p+ + +
      T259  644 ---SLHFEPSKINI    654
```
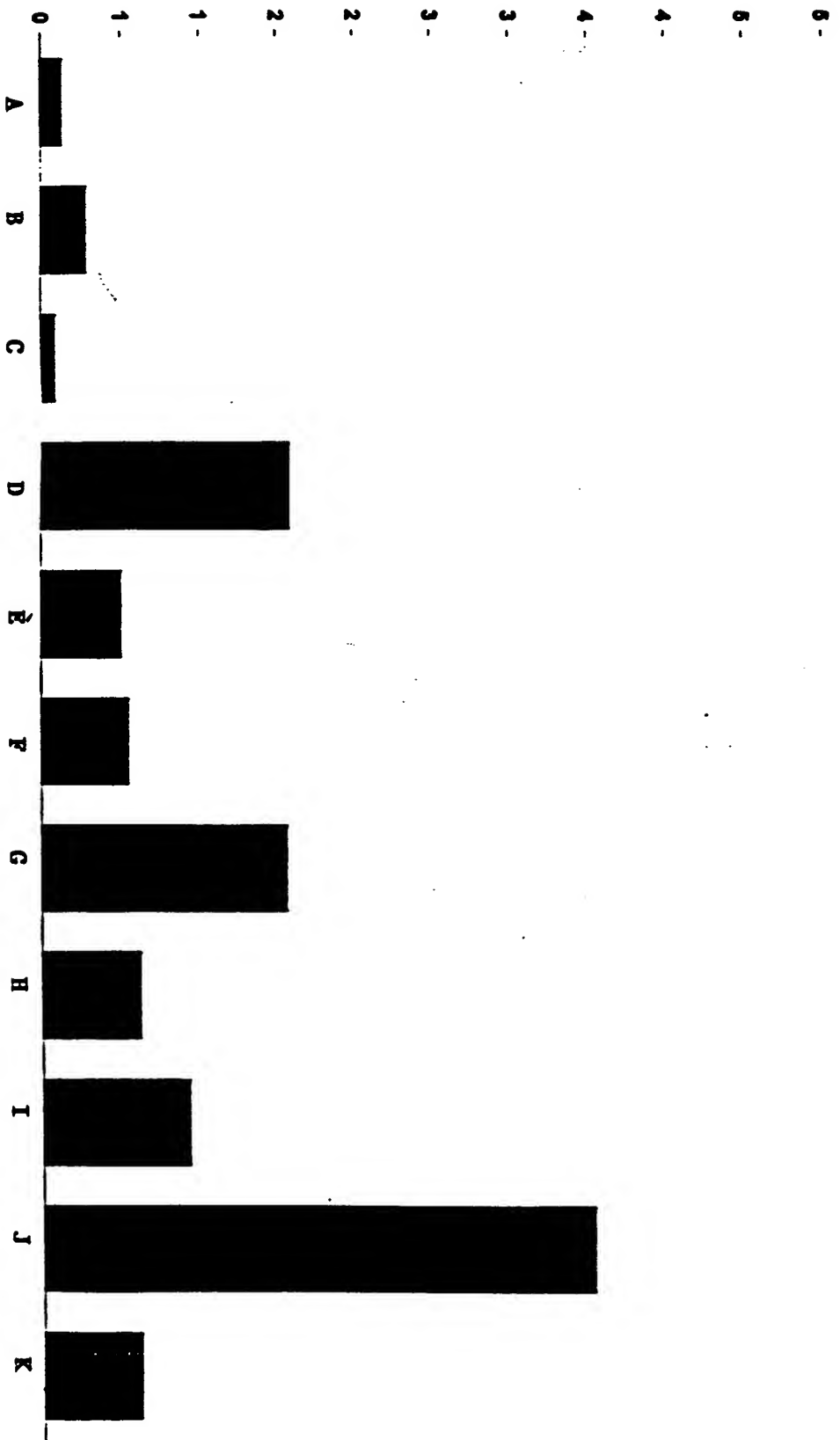
Figure 9E

Figure 10A

Relative Expression (NTC used as reference sample)

Figure 10B